APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Aug. 2000, p. 3310–3329 Vol. 66, No. 8
0099-2240/00/$04.00+0

A Genomic Sample Sequence of the Entomopathogenic Bacterium *Photorhabdus luminescens* W14: Potential Implications for Virulence

RICHARD H. FFRENCH-CONSTANT,[1]* NICHOLAS WATERFIELD,[1] VALERIE BURLAND,[2] NICOLE T. PERNA,[2] PHILLIP J. DABORN,[1] DAVID BOWEN,[3] AND FREDERICK R. BLATTNER[2]

Department of Biology and Biochemistry, University of Bath, Bath, BA2 7AY, United Kingdom,[1] and Laboratory of Genetics,[2] and Department of Entomology,[3] University of Wisconsin-Madison, Madison, Wisconsin 53706

Received 6 March 2000/Accepted 25 May 2000

***Photorhabdus luminescens* is a pathogenic bacterium that lives in the guts of insect-pathogenic nematodes. After invasion of an insect host by a nematode, bacteria are released from the nematode gut and help kill the insect, in which both the bacteria and the nematodes subsequently replicate. However, the bacterial virulence factors associated with this "symbiosis of pathogens" remain largely obscure. In order to identify genes encoding potential virulence factors, we performed ~2,000 random sequencing reads from a *P. luminescens* W14 genomic library. We then compared the sequences obtained to sequences in existing gene databases and to the *Escherichia coli* K-12 genome sequence. Here we describe the different classes of potential virulence factors found. These factors include genes that putatively encode Tc insecticidal toxin complexes, Rtx-like toxins, proteases and lipases, colicin and pyocins, and various antibiotics. They also include a diverse array of secretion (e.g., type III), iron uptake, and lipopolysaccharide production systems. We speculate on the potential functions of each of these gene classes in insect infection and also examine the extent to which the invertebrate pathogen *P. luminescens* shares potential antivertebrate virulence factors. The implications for understanding both the biology of this insect pathogen and links between the evolution of vertebrate virulence factors and the evolution of invertebrate virulence factors are discussed.**

Photorhabdus luminescens is an insect-pathogenic gram-negative proteobacterium that forms a "symbiosis of pathogens" with insect-pathogenic nematodes (52). In this symbiosis the bacteria are carried in the guts of entomopathogenic nematodes belonging to the family Heterorhabditidae (members of a different group of bacteria, *Xenorhabdus* spp., are carried in the guts of members of a different group of nematodes, the Steinernematidae). Upon invasion of an insect host by a nematode, the bacteria are released from the gut directly into the open blood circulatory system of the insect, the hemocoel (52). Here the bacteria are thought to release a wide variety of potential virulence factors, including high-molecular-weight toxin complexes (Tc), lipopolysaccharide (LPS), proteases, lipases, and a range of different antibiotics (52). Inferences concerning the involvement of these factors in killing of the insect or in overcoming the insect immune system, however, often result merely from documentation of secretion of the factors into bacterial culture supernatants. Studies examining the precise role of virulence factors during the infection process in insects have not been performed, and studies of *Photorhabdus* mutants are rare. As a prelude to genetic analysis of potential virulence factors in *P. luminescens*, we were interested in obtaining a sample sequence of strain W14 in order to document the classes of genes present and to begin to design suitable experiments for analysis of the genes based on a likely idea of their functions.

The relative advantages of sample sequence analysis versus full-scale analysis of a finished bacterial genome have been discussed elsewhere (119). However, there are several points relevant to the current discussion, as discussed briefly below. First, a sample sequence can be completed at a fraction of the cost of completion of a full genome. Second, a surprisingly high percentage of the genome can be captured even with a 1× sample sequence. Given the current uncertainty concerning the exact genome size of *P. luminescens*, the percent coverage obtained in this study is hard to estimate; however, McClelland and Wilson (119) suggested that a 1× genome equivalent for the 4.78-Mbp *Salmonella typhi* genome would require only 12,000 reads of 400 bases. Such coverage would ensure that almost every cistron was represented in the sample sequence. The 2,000 reads reported here obviously do not give this level of coverage, but, as shown below, even the limited sample sequence obtained revealed ample evidence concerning the types of virulence systems that *P. luminescens* may employ in its complex life cycle.

Although few potential *P. luminescens* virulence factors have been examined in detail (either biochemically or genetically), we can attempt to predict the likely role of bacterial virulence systems in killing an insect, in overcoming an insect immune system, or in facilitating bacterial and/or nematode growth. It is thought that once *P. luminescens* is released from the nematode gut into the insect hemocoel, it plays multiple roles in helping the nematode overcome its host (52). To do this, the bacteria need to overcome both the cellular (hemocytic) and peptide-mediated (antibacterial polypeptide) components of the insect immune system. Furthermore, the bacteria stop the insect from feeding and probably render its tissues suitable for consumption by both the bacteria and the nematodes. Anti-insect virulence mechanisms might, therefore, include, but not be limited to, toxins active against the insect gut and/or hemocytes and enzymes (such as proteases) capable of both degrading insect tissue and disabling the antibacterial peptides also associated with the insect immune system. Equally important

* Corresponding author. Mailing address: Department of Biology and Biochemistry, South Building, University of Bath, Bath, BA2 7AY, United Kingdom. Phone: 44 1225 826261. Fax: 44 1225 826779. E-mail: bssrfc@bath.ac.uk.

in its role in overcoming an insect host, *P. luminescens* must ensure that the insect cadaver does not act as a breeding ground for opportunistic soil bacteria, fungi, and/or other species of nematodes. We might, therefore, expect *P. luminescens* to secrete a wide range of antimicrobial, antifungal, and nematicidal compounds, as previously documented by other workers (52). The aim of the present study was, therefore, to identify genes that encode likely virulence factors as a prelude to a functional analysis of the genes via targeted knockout and assay of the resulting mutants in the host infection process. Not only should such an analysis allow us to elucidate how the virulence factors act on the insect, but the gene sequences may also provide an indication of the evolution and potential origins of the virulence factors.

MATERIALS AND METHODS

Genomic library construction and sequencing. Genomic DNA from *P. luminescens* W14 was size selected to obtain 1- to 2-kb fragments and then cloned into M13 Janus as previously described (28, 115). DNA templates were purified from library clones and sequenced by using dye terminator-labeled fluorescent cycle sequencing (model ABI377 automated sequencer; Applied Biosystems Division, Perkin-Elmer). Single sequencing reads (average length, ~400 bp) were obtained for one end of 2,122 clones. Sequences were truncated to exclude the phage arms and multiple cloning site and were then submitted to the BLASTX servers at the National Center for Biotechnology Information. Clones giving hits to either Tc-, protease-, or Rtx-like-encoding genes were then sequenced from the other end or "flipped."

Comparison with *Escherichia coli* K-12. Trimmed (vector removed and high-quality trim with SeqManII) *P. luminescens* sequence reads were searched against the DNA and protein sequences of *E. coli* MG1655 by using BLASTN and BLASTX with a local server. The output was parsed and sorted to give three subsets of data with different levels of identity. No alignment length criteria were imposed on the output. The results, therefore, included short alignments and multiple hits for many sequences, all of which were legitimate similarities.

Nucleotide sequence accession number. The nucleotide sequence determined in this study has been deposited in the DDBJ/EMBL/GenBank database under accession no. AQ989457–AQ991805.

RESULTS AND DISCUSSION

Comparison with *E. coli* K-12. The results of the comparison of *P. luminescens* sequences with *E. coli* K-12 sequences are shown in Fig. 1. Even though the number of query sequences was relatively small (<0.5× genome coverage), we clearly observed that there is significant conservation of sequences, particularly protein sequences, between the two genomes, which is consistent with the relatively close phylogenetic relationship of the two organisms (both are members of the family *Enterobacteriaceae*). Regions of the genome conserved in *Escherichia* and *Photorhabdus* strains begin to define the components of the putative ancestral chromosome of members of the gamma subdivision of the class *Proteobacteria*. The large excess of hits at the protein level compared to the DNA level at all stringencies suggests that the divergence between orthologous sequences is sufficient to obscure true matches at the nucleotide level. Even the number of protein hits changed extensively as we varied the criteria for a significant match. Thus, we were reluctant to choose an arbitrary cutoff for determining orthology (as has been done for other sample sequence comparisons) and instead describe three different levels of stringency below. This form of presentation provides not only a sense of the absolute number of sequences that are similar but also a sense of the strength of the similarities. It should be noted that although the hits are distributed, albeit unevenly, around the K-12 map, this does not necessarily indicate that there is colinearity, and indeed the sizes of the two genomes are probably different, which could account for some of the gaps and sparse regions in Fig. 1. In all, 1,133 W14 clones exhibited no significant matches in even the lowest-stringency analysis ($E < E^{-05}$), and from this we inferred that approximately 53% (1,133 of the 2,122 clones examined) of the *P. luminescens* genome is clearly distinct from the genome of *E. coli* K-12.

Old and new *toxin complex* (*tc*) loci. We previously cloned and sequenced four Tc-encoding loci, *tca*, *tcb*, *tcc*, and *tcd*, from *P. luminescens* W14 (22); each of these loci encodes a different high-molecular-weight insecticidal Tc (Tca, Tcb, Tcc, and Tcd, respectively). The Tc proteins are secreted into the supernatant by *P. luminescens* grown in liquid culture (23). Despite the fact that the Tc toxins exhibit both oral and injectable activity against a range of insects (22, 23), their precise role as potential virulence factors in the infection process remains to be determined. However, one of the complexes, Tca, has highly specific histopathological effects on the lepidopteran midgut (18), suggesting that Tca proteins may be used by the bacterium to destroy the insect midgut and effectively stop feeding. In the sample sequence analysis, BLASTX searches gave 19 hits for the four known *tc* loci (22), but 27 additional sequences (Table 1) were also identified that could not be ascribed to the previously identified *tc* loci after careful examination of the sequence chromatographs (Fig. 2A). This suggests that there are other *tc*-like loci in the *P. luminescens* W14 genome in addition to those already reported. The matches with new *tc*-like loci were classified as *tca*-like (3 hits), *tcc*-like (13 hits), or *tcb*/*tcd*-like (11 hits; *tcb* and *tcd* are close homologs of one another).

The hypothesis that there are more than four *tc* loci in the W14 genome was confirmed by several other lines of evidence. First, extended sequencing of DNA surrounding the *tcdA* locus revealed not only the presence of a second open reading frame (ORF) immediately downstream of *tcdA* (designated *tcdB*) but also the presence of a second *tccC*-like locus further downstream from the *tcdAB* locus (unpublished results). As suggested by the sample sequence, this proves that there are at least two copies of *tccC* in the W14 genome. Second, sequencing of the opposite ends of flipped *tc*-containing clones showed that some of the new *tc*-like loci occupy novel genomic positions beyond the positions established for the four known loci. For example, clone 02349 is a *tccA*-like sequence whose flip (clone 02349f) is a *lon* protease, and clone 01515 is a *tccB*-like sequence whose flip is an exochitinase. Clone 00763 contains a *tccC*-like sequence which forms a contig with three other clones (00763f, 00339, and 02380), which also contain a *yfiP*-encoded lipase. Finally and perhaps most interestingly, one sequence (00357) contains both *tccC*-like and *tcaC*-like sequences but has phage sequences inserted between them (Fig. 2B shows the implied genomic organization of this contig). The abundance and potential implications of phagelike sequences in the *P. luminescens* W14 genome are discussed below. However, together, the sequence and inferred position data provide firm evidence that additional *tc* loci are present in the W14 genome. The implications for the potentially increased variety of encoded insecticidal Tc toxins remain unclear.

Antibiotics and antibiotic resistance. Having destroyed the insect gut, presumably by using the *tc*-encoded Tc (18), *P. luminescens* must then defend the insect cadaver from a wide range of other colonizing organisms, such as bacteria (including other strains of *P. luminescens*), fungi, and/or nematodes. It seems reasonable to assume that during this process *P. luminescens* W14 deploys a range of antimicrobial agents, such as antibiotics and antifungal agents, as documented for other *Photorhabdus* strains and also *Xenorhabdus* strains (52), in order to maintain a bacterial monoculture in the insect cadaver. Thus, in the sample sequence one of the largest classes of hits was hits for polyketide synthetase-like genes. This class of genes is responsible for nonribosomal synthesis of a diverse array of compounds involved in processes ranging from fatty

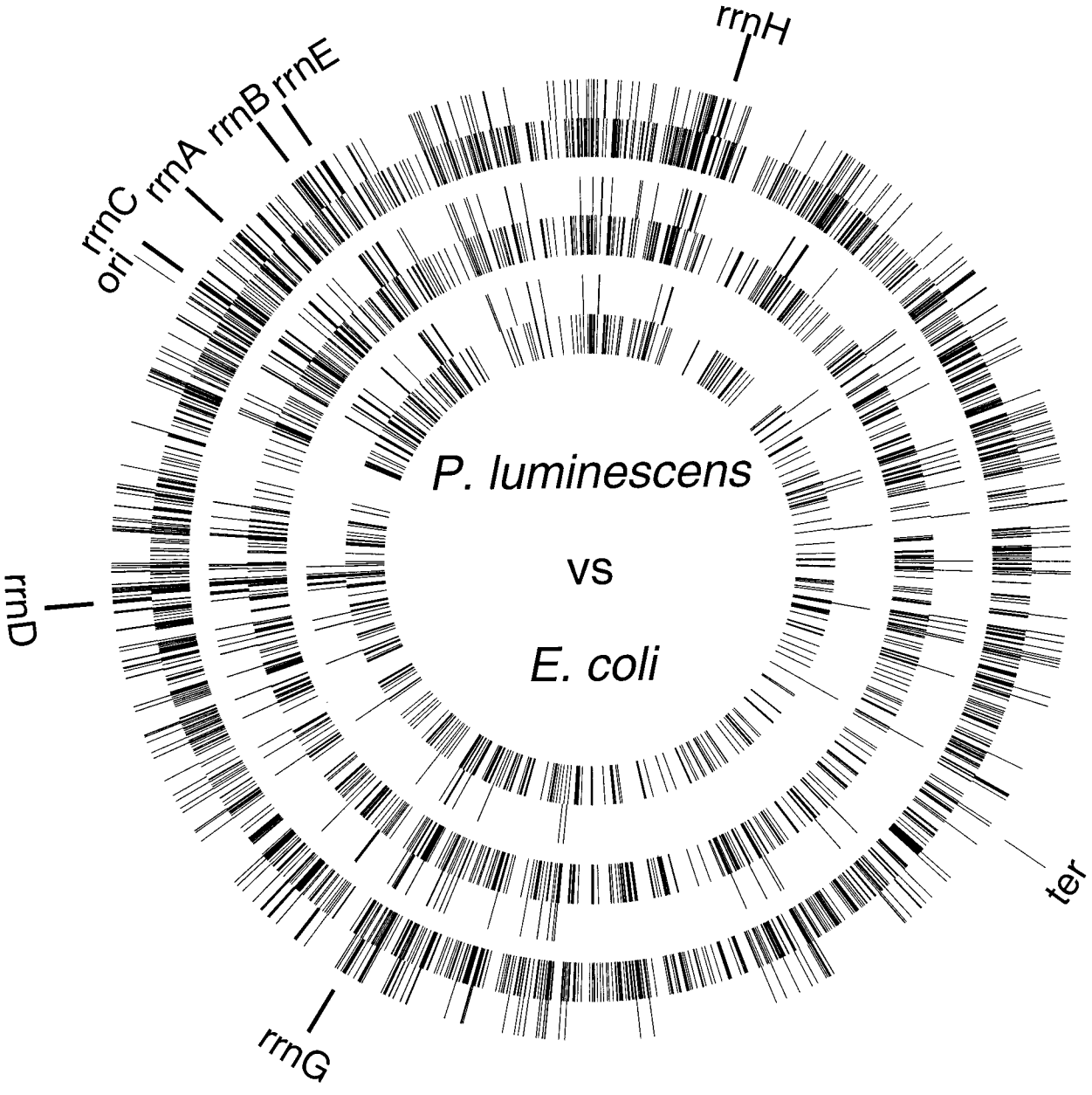

FIG. 1. Graphic display of sample sequence similarities to *E. coli* K-12 nucleotide and protein sequences, generated from a BLAST search of *P. luminescens* sequences performed with K-12. The three concentric sets of data show nucleotide (outer ring) and protein (inner ring) hits plotted at the coordinates of the K-12 target. From the outside, the three data sets show hits with BLAST expected value (*E*) limits of $<10e^{-05}$ (2,765 protein and 729 nucleotide hits), $<10e^{-20}$ (1,227 protein and 376 nucleotide hits), and $<10e^{-40}$ (664 protein and 234 nucleotide hits), respectively. The positions of the genetic markers *ori*, *ter*, and *rrnA* through *rrnH* are shown as landmarks to orient the circle. The figure was generated by using the program Genescene (DNASTAR).

acid synthesis to antibiotic production (including production of inhibitors of eukaryotic protein phosphatases [41]). Even if we took into account the effects of the large sizes of some of the polyketide synthetase loci (up to 28 kb of repeated subunits), these classes of hits were still some of the predominant classes of hits in the sample sequence, accounting for 3.7% (80 hits) of the total sequences. Of the matches with polyketide synthetase-like sequences, 31 were with a syringomycin synthetase from *Pseudomonas syringae* pv. syringae (Table 2). Interestingly, the syringomycin synthetase gene cluster is thought to provide a link between prokaryotic and eukaryotic peptide synthetases (62), while syringomycin itself has a wide range of antibacterial and antifungal properties. Deployment of similar antibiotics by *P. luminescens* W14 may, therefore, help maintain a bacterial monoculture in an insect cadaver. Furthermore, it is interesting to note that *P. luminescens* also contains a sequence that is similar to tolaasin (another lipodepsipeptide), which is used for self-protection in *Pseudomonas tolaasii*, which implies that W14 may employ this peptidoglycan-associated lipoprotein in self-protection against its own antibiotics. In addition to potentially deploying broad-spectrum antibiotics to repel other organisms that might colonize the insect cadaver, strain W14 also contains sequences similar to colicin activity proteins (CeaAB), colicin transport proteins (BtuB),

TABLE 1. Hits to predicted products of known *tc* loci and putative new *tc*-like loci from *P. luminescens*[a]

Gene function	Accession no.	No. of Hits[b]	BLASTX *E* value(s)	Clone(s)	Reference
Known *tc* loci					
TcaABCZ	AF046867	3 (1)	3e-11 to 1e-93	01421, 01421f, 02319	22
TcbA	AF04757	5 (0)	2e-25 to 1e-102	00513, 00753, 00753f, 00949, 01893	22
TccABCZ	AF47028	9 (3)	1e-53 to 8e-91	00707, 00707f, 01179, 01478, 01478f, 02015, 02197, 002197f, 02327	22
TcdA	AF188483	2 (0)	3e-34 to 6e-84	00617, 01839, 02280	Unp.[c]
Putative new *tc*-like loci					
TcaA-like	AF046867	1 (0)	2.6	01461	22
TcaC-like	AF046867	2 (0)	6e-14, 1e-14	00357, 01661[d]	22
TcbA- and TcdA-like	AF04757	11 (5)	7e-04 to 2e-51	00508, 00508f, 00598, 00598f, 00878, 01303, 01303f, 01508, 01744, 01939, 02105, 02189, 02507, 02507f, 02474, 02474f	22
TccA-like	AF047028	5 (1)	0.01 to 2e-37	00093, 00093f, 01817, 01483, 02281, 02349	22
TccB-like	AF047028	3 (0)	9e-16 to 4e-34	01932, 01515, 01932	22
TccC-like	AF047028	5 (1)	6e-07 to 3e-79	00357f, 00763, 00869, 01403, 01498, 02049	22

[a] See Fig. 2 for genomic locations.
[b] Number of sequence hits. The numbers in parentheses are the numbers of flipped sequences (indicated by the suffix f in the clone designations).
[c] Unp., unpublished data.
[d] Clone 00357 represents a sequence containing two different ORFs.

and pyocin immunity proteins (S3). A sequence similar to colicin lysis protein was not found, although there was a match with a similar VlyS lysis protein S from lambda phage (BLASTX *E* value, 2e-16). Although the role of the colicin- or pyocin-like sequences in *P. luminescens* remains to be determined, they may be used to produce toxins and antitoxins designed to kill non-self bacteria.

In addition to genes for specific mechanisms for antibiotic production and self-protection, the W14 genome contains numerous sequences that exhibit homology to genes for other antibiotic resistance mechanisms. These sequences include genes involved in resistance to penicillin (penicillinase and penicillin-binding protein), bicyclomycin, and a range of other antibiotics (tetracycline, rifampin, and kasugamycin) via a variety of different mechanisms (Table 3). Most notable in this respect are the large number of sequences that exhibit homology to genes for different multiple-drug-like export systems, including Emr-like and Mdl-like systems that export drugs ranging from chloramphenicol to acriflavin. These multiple-drug export systems are also very similar to the hemolysin B export systems, as discussed below, and begin to describe a large family of exportlike genes in the *P. luminescens* genome. Also present are sequences similar to cation resistance genes in other enteropathogenic bacteria, notably sequences that encode resistance to tellurite (TelA) in *E. coli* plasmid RK2 (Table 3).

Rtx-like homologs. Another large class of database matches comprises sequences similar to both Rtx-like and hemolysin A-like toxins and their associated export systems (Table 4). The RTX (repeats in toxin) toxins are cytolytic toxins that are virulence factors in many pathogenic gram-negative bacteria (182). The RTX elements of other gram-negative bacteria share certain aspects of genomic organization, including the presence of three elements: an exported protein (RtxA-like), an ATP-binding cassette ABC protein (RtxB-like), and a membrane fusion protein (RtxD-like). Figure 3 shows the sequences similar to each of these elements alongside the loci to which they are most similar as determined by BLASTX searches. This figure shows that the *Photorhabdus* sample sequence contains sequences similar to the sequences of RtxA and RtxB of *Vibrio cholerae*, ShlA and ShlB of *Serratia marcescens*, EthA and EthB of *Erwinia tarda*, and HecA and HecB of *Erwinia chrysanthemi*. We also discerned sequences similar to both HlyB and CvaA/CvaB of *E. coli*, which are involved in hemolysin secretion and colicin V secretion, respectively. Notably, even if we took into account the large predicted ORF size (size of *rtxA*, ~12 kb), there were still 24 hits with RtxA-like sequences alone, suggesting that more than one locus may be present. Furthermore, BLASTX *E* values were highly significant (e-25 to e-78), suggesting that there is a high level of amino acid conservation. We also observed that there is a sequence similar to TolC which is unlinked but is required both for hemolysin export and for colicin V export (58). We can only speculate as to the number of loci that these sequences correspond to and to the likely role of the encoded toxins in *P. luminescens* infection. However, given the propensity of Rtx-like toxins to attack host phagocytes (182), we postulate that the sequences may be important in attacking the insect cellular immune system, the hemocytes. This hypothesis could be tested by deleting the toxin loci or their export machinery and examining the infection process in their presence and absence.

In addition to an RtxA-like export system, we also found evidence of other Rtx-like export systems, including an Rtx-like metalloprotease and its accompanying export machinery. The Rtx-like metalloprotease itself is similar to PrtA-encoded metalloproteinase A of *E. chrysanthemi*, while its associated export machinery is similar to the LipBCD-like ABC transporter of *S. marcescens*, which also exports a protease. Between the protease and its associated ABC transporter there is a small protease inhibitor (as confirmed by our extended sequencing of the operon). This genomic organization of an Rtx-like metalloprotease and its associated LipBCD-like transporter shows that *P. luminescens* uses different combinations of Rtx-like genes to export virulence factors and stresses the potential importance of these systems for anti-insect virulence.

Other putative virulence factors. In addition to the specific Tc-like and Rtx-like toxins discussed above, we also identified a wide range of other sequences related to a diverse array of genes that are potentially involved in infection and virulence. These genes include genes that encode factors involved in bioluminescence, other proteases, lipases, hemmaglutinins, chitinases, and other toxins, such as non-Rtx hemolysins and ADP-ribosyltransferases (Table 5). They also include genes involved in two-component sensor systems that have previously been implicated in regulation of virulence both in *Photorhabdus* strains and in other bacteria.

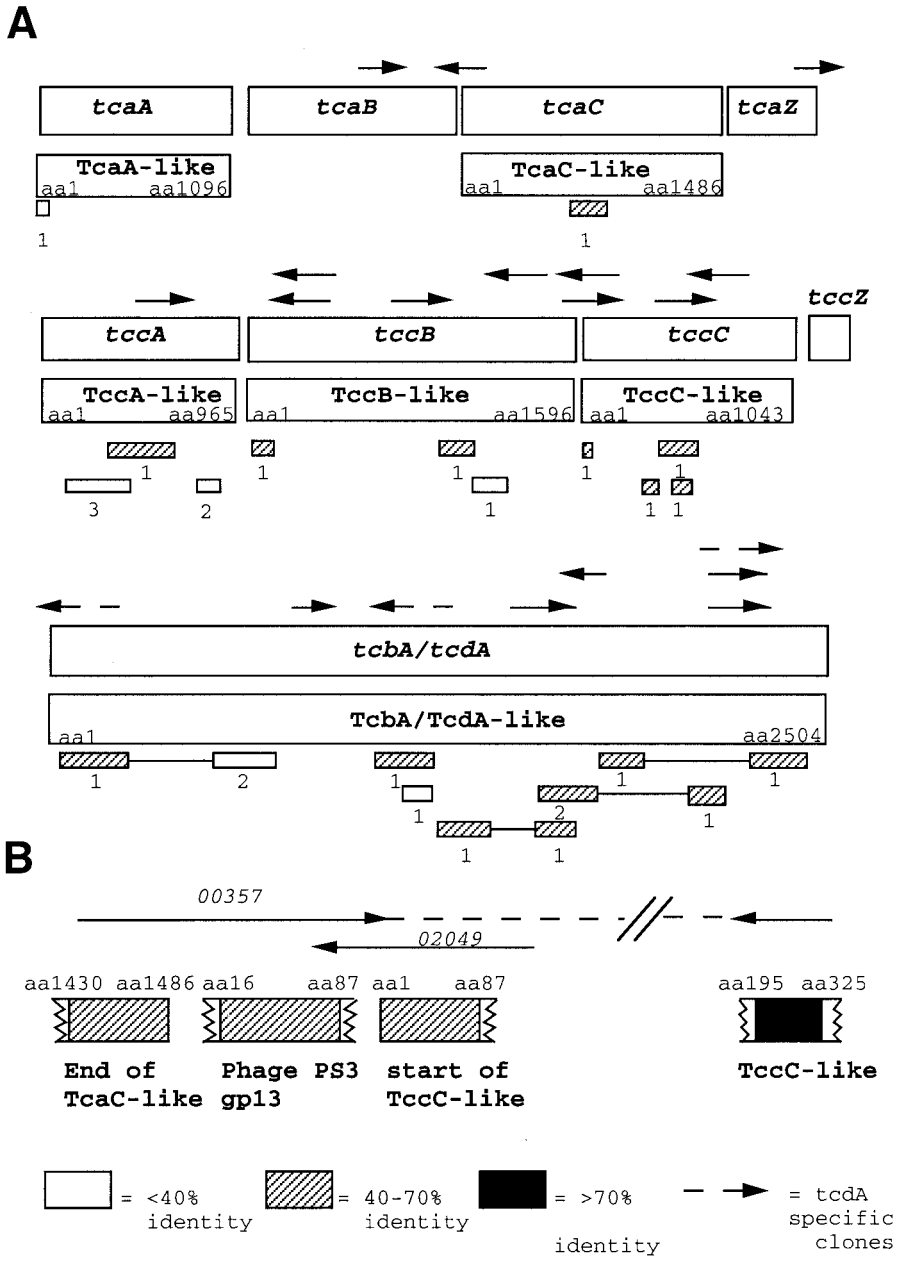

FIG. 2. Diagrams showing the relative locations of hits to known *tc* loci and new *tc*-like loci. (A) Locations of individual sequencing reads (arrows above the diagrams) and their associated contigs and BLASTX matches (boxes below the diagrams). Note that the predicted amino acid sequences for *tcb* and *tcd* are sufficiently similar that we could not distinguish matches with either locus. (B) One example of a difference in genomic organization of a new *tc*-like locus inferred from a small contig and adjacent flipped sequence. Note that two TccC-like BLASTX matches are located next to a TcaC-like ORF with a phage remnant in between (see text). aa, amino acid.

Bioluminescence (from which *P. luminescens* obtained its specific epithet) occurs shortly after bacteria invade an insect, but its biological role is unclear. The genes that encode the luciferase beta subunit and NAD(P)H-flavin reductase, which reduces flavin mononucleotide for bioluminescence, have previously been cloned, and hits to these genes were highly significant (BLASTX *E* values, 1e-39 to 3e-66), which again confirmed the quality and coverage of the sample sequence. Compared to other classes of potential virulence factors, we found several sequences similar to sequences of non-RTX-like proteases and lipases, which have previously been implicated in virulence. One of these, triaglycerol lipase 1, has been cloned previously, and hits to this sequence were highly significant (BLASTX *E* values, 4e-13 to 9e-86). Other proteins, like the Lys-X cysteine protease of *Porphyromonas gingivalis*, have been implicated in virulence during soft-tissue infections (109). Another subclass of hits in this category are hits to matrix metalloprotease-like sequences. These are interesting because one of the proteins, limunectin (from the horseshoe crab, *Limulus* sp.), binds bacterial cells, fixed amebocytes, and extracellular matrix molecules (113). If *Photorhabdus* cells do indeed make a similar protein, the protein may play some role in bacterial aggregation, previously termed nodulation (52), or in attachment to host cells. Other hits to proteins potentially

TABLE 2. Hits to polyketide synthetase-like proteins, colicin, and pyocins

Gene function	Organism	Accession no.	No. of hits[a]	BLASTX *E* value(s)	Clone(s)	Reference
Polyketide synthetases: syringomycin synthetase-like						
Syringomycin synthetase (*Pseudomonas*)	*P. syringae* pv. syringae	AF047828	31 (22)	2e-09 to 3e-54	00006, 00006f, 00033, 00033f2, 00104, 00134, 00134f, 00033, 00033f, 00104, 00134, 00134f, 00554, 00554f, 00564, 00564f, 00967, 00976, 00967f, 01218, 01247, 01458, 01458f, 01986, 00015, 00037, 00037f, 00380, 00380f, 00665, 00665f, 01057, 01190, 01190f, 01258, 01385, 01385f, 01513, 01519, 01519f, 01762, 01901, 01901f, 02274, 01973f, 00649, 00502f, 00060, 00466f, 00498f, 00533f, 00379, 00772f, 01029, 01170, 01258f, 01633, 01946	62
Other polyketide synthetase-like proteins						
BacA, bacitracin synthetase 1	*Bacillus licheniformis*	AF007865	2 (0)	1e-06 to 3e-29	01477, 00946	97
BacC, bacitracin synthetase 3	*B. licheniformis*	AF007865	7 (2)	4e-06 to 3e-29	00058, 00058f, 00996f, 01242, 01614, 01782, 01783	97
Saframycin Mx1 synthetase A	*Myxococcus xanthus*	U24657	1 (0)	1e-04	02574	143
Saframycin Mx1 synthetase B	*M. xanthus*	U24657	5 (2)	1e-11 to 4e-29	00060f, 004466, 01119, 01119f, 02368	143
Pristinamycin I synthetase	*Streptomyces pristinaespiralis*	X98690	3 (1)	4e-04 to 2e-24	00502, 00649f, 00772	39
LicA, lichenysin synthetase A	*B. licheniformis*	U95370	3	4e-09 to 1e-26	00750, 01729, 01171	96
LicB, lichenysin synthetase B	*B. licheniformis*	U95370	2 (2)	4e-09, 2e-14	00750f, 00946f	96
Tyrocidine synthetases 1, 2, and 3	*Bacillus brevis*	AF004835	4 (2)	5e-13 to 3e-16	00641, 00738, 01242f, 02574f	126
PksK, polyketide synthetase	*Bacillus subtilis*	P40803	2 (1)	0.4, 7e-35	00269f, 01717	3
PksL polyketide synthetase	*B. subtilis*	P40803	1	0.04	02466	3
PksR, polyketide synthetase	*B. subtilis*	P40803	1	4e-04	01729f	3
Surfactin synthetase subunit 2	*B. subtilis*	Q04747	3 (1)	4e-10 to 7e-36	00104f, 00738f, 02068	20
Peptide synthetase-like	*B. subtilis*	AF087452	3 (0)	3e-07 to 7e-08	00153, 00136, 01474	Unp.[b]
Lysobactin synthetase	*Lysobacter* sp.	X96558	3 (1)	6e-17 to 6e-23	00533, 00641f, 00996	17
Daptomycin synthetase-like	*Streptomyces roseosporus*	AF021263	2 (0)	1.1, 4e-34	02364, 01805	121
Polyketide synthetase 6-like	*Mycobacterium tuberculosis*	Z84725	2 (0)	1e-12 to 3e-14	01953, 02416	34
Danorubicin-like	*Streptomyces peucetius*	L35560	1 (0)	8e-09	02548	170
Pyoverdine synthetase D-like	*P. aeruginosa*	S53999	2 (0)	9e-07, 6e-39	00812, 01517	122
Gramicidin S synthetase 2	*B. brevis*	JX0340	1 (0)	9e-31	00097	152
Microcystin synthetase B	*Microcystin aeruginosa*	U97078	1 (0)	2e-25	01315	41
Saccharopolyspora PKS	*Saccharopolyspora hirsuta*	S35197	1 (0)	2e-14	01493	104
Self-protection to lipodepsipeptides						
Pal, peptidoglycan-associated lipoprotein	*E. coli*	P07176	1 (0)	3e-38	01801	32
Tolassin self-protection	*P. tolaasii*	U16024		1e-09	(Second match)	Unp.
Colicins and pycocins						
CeaA, colicin A	*Citrobacter freundii* plasmid	P04480	1 (0)	6e-24	01036	127
CeaB, colicin activity protein	*E. coli* pColE2	P04419	1 (0)	4e-34	00284	116
Pyocin S3 immunity protein	*P. aeruginosa* P12	B56394	3 (1)	4e-06 to 2e-15	01787, 02024, 02024f	44
BtuB, transport of E colicins	*C. freundii*	Y09059	1 (0)	1e-59	00358	Unp.
YebA, hypothetical lysostaphin	*E. coli*	P24204	1 (0)	2e-60	07734	76

[a] Number of sequence hits. The numbers in parentheses are the numbers of flipped sequences (indicated by the suffix f in the clone designations).
[b] Unp., unpublished data.

TABLE 3. Hits to antibiotic and drug resistance-associated proteins

Gene function	Organism	Accession no.	No. of hits[a]	BLASTX *E* value(s)	Clone(s)	Reference
Penicillinase and penicillin-binding proteins						
BlaC, penicillinase	*Y. enterocolitica*	Q01166	2 (0)	3e-11, 5e-59	02059, 02237	157
PbpB, penicillin-binding protein 1B	*E. coli*	AE000124	1 (0)	3e-11	01997	56
PbpE, penicillin-binding protein 4	*Bacillus subtilis*	P32959	2 (0)	2e-07, 2e-10	00417, 02075	142
Bicyclomycin resistance proteins						
YnfM, bicyclomycin resistance	*E. coli*	D90801	2 (0)	3e-51, 2e-52	00114, 02573	Unp.[b]
Bcr, bicyclomycin resistance	*E. coli*	AE000308	1 (0)	3e-18	00511	19
Bcr, bicyclomycin resistance	*H. influenzae*	P45123	1 (0)	3e-13	02347	51
Other antibiotic resistance proteins						
TypA, GTPase	*E. coli*	AJ224871	1 (0)	1e-76	01716	47
LpxD, glucosamine *N*-acyltransferase (rifampin)	*Y. enterocolitica*	P32203	1 (0)	4e-29	00423	180
KsgA, dimethyladenosine transferase (kasugamycin)	*E. coli*	P06992	1 (0)	2e-74	00122	175
Multiple-drug-like efflux systems						
Streptomyces chloramphenicol resistance-like	*B. subtilis*	AB001488	2 (0)	3e-09, 3e-13	01467, 01540	Unp.
EnvD, protein D (acriflavin)	*E. coli*	D90846	3 (0)	3e-26 to 6e-71	01292, 01981, 02167	91
AcrE, acriflavin resistance-like	*Aquifex aeolicus*	AE000702	1 (0)	5e-06	01439	Unp.
Emr multiple-drug-resistance proteins						
EmrD, protein D	*E. coli*	P31442	4 (0)	3e-10 to 3e-56	00349, 01526, 01647, 01875	129
EmrY, protein Y	*E. coli*	P52600	1 (0)	3e-53	01983	19
Mdl multiple-drug protein						
Mdl, ATP binding	*E. coli*	U82664	2 (0)	4e-53, 5e-76	01456, 01960	Unp.
Other putative resistance-associated translocases						
YqjV, resistance protein-like	*B. subtilis*	P54559	2 (0)	1e-07, 1e-09	00530, 01766	Unp.
YfkF, resistance protein-like	*B. subtilis*	D83967	2 (0)	0.002 to 6e-04	01266, 02238	99
YgeD, resistance protein-like	*E. coli*	P39196	1 (0)	2e-50	00624	77
YieO, resistance protein-like	*E. coli*	P31474	1 (0)	4e-18	02452	29
YbhF, ATP binding	*E. coli*	P75776	2 (1)	3e-40, 3e-43	02378, 02377f	19
Ethidium bromide resistance						
E1 protein (putative chaperone)	*E. coli*	D90802	4 (0)	1e-08 to 3e-27	00009, 00030, 00077, 00666	1
Cation and solvent resistance-like						
YaaN, toxic cation resistance	*B. subtilis*	P37535	1 (0)	6e-41	02039	133
TelA, tellurite resistance	*E. coli* pRK2	Q52328		2e-29		60
OstA, organic solvent tolerance	*E. coli*	P31554	2 (0)	3e-68, 5e-70	00135, 00828	6
Ttg2F, toluene tolerance	*Pseudomonas putida*	AF106002	1 (0)	2e-09	01702	90

[a] Number of sequence hits. The numbers in parentheses are the numbers of flipped sequences (indicated by the suffix f in the clone designations).
[b] Unp., unpublished data.

involved in host cell binding included hits to several hemagglutinins. For example, deletion of the filamentous hemagglutinin locus in *Bordetella pertussis* results in loss of binding to ciliated eukaryotic host cells (148). As well as degrading host cells via proteolytic activity, much of the insect exoskeleton is composed of chitin. Therefore, hits to chitinases (*N*-acetyl-beta-glucosaminidase) probably indicate that there are several chitin-degrading enzymes, notably an enzyme similar to the *chB*-encoded chitinase of *S. marcescens* and a chitinase C-like product of an insect (*Glossina morsitans*) S endosymbiont (Table 5).

In addition to the Tc and Rtx-like toxins discussed above, W14 also appears to contain several non-Rtx hemolysins and other classes of toxins. One of the non-Rtx hemolysins, hemocyte erythrocyte lysis protein 2 from *Prevotella intermedia*, is notable in that searches of DNA and protein databases have not previously revealed any significant homologies (14); the *Photorhabdus* homology reported here is, therefore, possibly such a match. Other classes of toxins include two ADP-ribosyltransferases and cytotoxic necrotizing factor 2, all of which are cytotoxins. The ADP-ribosyltransferases are like exotoxin A from *Clostridium difficile* and *Pseudomonas aeruginosa*. Although the BLASTX *E* values for these hits are of low significance (0.05 and 0.8), the narrow ranges of homology values indicate that the levels of predicted amino acid identity are high (44% [17 of 38 residues] and 32% [24 of 73 residues], respectively). Cytotoxic necrotizing factor 2 from *E. coli* acts on the small GTP-binding protein Rho involved in actin cytoskeleton assembly and causes stress fiber formation in target cells (137). It is also interesting that there was a hit to halovibrin

TABLE 4. Rtx-like operon homologs, including proteins exported by these operons and their accompanying export machinery and activating and modulating proteins[a]

Gene function	Organism	Accession no.	No. of hits[b]	BLASTX *E* value(s)[c]	Clone(s)	Reference
Exported proteins						
RtxA-like	*V. cholerae*	AF119150	17 (2)	2e-25 to 8e-78	00028, 00397, 00486, 00727, 00987, 00987f, 01071, 01071f, 01700, 01342, 01389, 01452, 01551, 02082, 02096, 02396, 02287, 02505, 02505f, 00727f, 02082f, 02333	112
EthA, hemolysin	*E. tarda*	D89876	4 (1)	0.003 to 4e-29	01557f, 01727, 01896, 02259, 01618, 00774	67
ShlA, hemolysin A	*S. marcescens*	P15320	2 (0)	1.8 (50), 1e-42	00369, 02511f, 00489f, 00369f	141
HmpA, hemolysin A	*P. mirabilis*	P16466	1 (1)	0.12 (30)	01539	174
PrtA, metalloproteinase A	*E. chrysanthemi*	JN0891	1 (1)	2e-47	01175f	24
Prt1, metalloproteinase A	*Erwinia carotovora*	Q99132	1 (1)	3e-18	00148	100
Activator or modulating proteins						
ShlB-like, hemolysin secretion	*S. marcescens*	P15321	6 (0)	3e-15 to 4e-54	00489, 01578, 02066, 02427, 02511, 02512	141
HecB-like, hemolysin secretion	*E. chrysanthemi*	L39897	5 (0)	1e-07 to 5e-13	00075, 00149, 00200, 01845, 02556, 00904f, 00149f, 01657	12
Protease inhibitor	*E. chrysanthemi*	AF071511	2 (0)	5.2 (47), 1e-15	00886, 01175[d]	106
Secretion functions: ATP-binding cassette proteins						
RtxB-like	*V. cholerae*	AF119150	2 (0)	5e-56, 2e-77	00848, 02333, 00848f	112
HlyB-like, hemolysin secretion	*E. coli*	P08716	2 (0)	2e-33, 4e-57	00909, 01791, 00909f	48
LipB, protease transporter	*S. marcescens*	D49826	7 (1)	6e-10 to 4e-65	00179, 00264, 00337, 01175*, 01512, 01512f, 02149	2
CvaB, colicin V secretion	*E. coli*	P22520	2 (1)	1e-09, 5e-13	00916, 00949f	58
Membrane fusion proteins						
CvaA, colicin V secretion	*E. coli*	P22519	2 (0)	1e-15, 2e-13	01283, 01780	58
	V. cholerae	AF119150		1e-15, 2e-08		112
LipC, protease transporter	*S. marcescens*	D49826	2 (1)	5e-37, 2e-50	00179f, 02010	2
Outer membrane proteins						
TolC, outer membrane protein	*Salmonella enteritidis*	Q54001	1 (0)	3e-61	00783	156
TolA, outer membrane protein	*E. coli*	P19934	1 (0)	3e-15	01818	108
LipD-like (PrtF, TolC)	*E. chrysanthemi*	P23598	1 (0)	0.8 (27)	00181	107
Hemolysin coregulated protein						
Hcp, 28-kDa secreted protein	*V. cholerae*	S911006	1 (0)	2e-09	01146	184

[a] See Fig. 3 for putative genomic organizations.
[b] Number of sequence hits. The numbers in parentheses are the numbers of flipped sequences (indicated by the suffix f in the clone designations).
[c] The numbers in parentheses are percentages.
[d] Clone 01175 represents a sequence containing two different ORFs.

from *Vibrio fisheri*, which is a member of a novel class of ADP ribosyltransferases with no significant sequence homology to other ADP ribosyltransferases (147). In relation to potential ADP ribosyltransferase regulation, the sample sequence had a highly significant (BLASTX *E* value, 6e-50) hit to ExsA, the exoenzyme S synthesis regulatory protein (53). Exoenzyme S is another ADP ribosyltransferase that is distinct from exotoxin A and is secreted by *P. aeruginosa*, and ExsA is an AraC-like transcriptional regulator of its production. This sequence is also similar (BLASTX *E* value, 2e-44) to the VirF virulence regulon transcriptional regulator which controls the *yop* regulon (see below). Finally, with regard to other non-Tc toxins, there were two low-scoring hits to the delta-endotoxins from *Bacillus thuringiensis*, whose significance is uncertain.

Hits on other potential virulence factors included matches to Vac, Vap, and Kic-like proteins. There were hits on VacB from both *E. coli* and *Haemophilus influenzae*. Disruption of this gene in enteroinvasive *E. coli* results in reduced expression of virulence phenotypes, suggesting that it is necessary for full expression of virulence (172). We also found sequences similar to both VapD and VapZ from *Dichelobacter nodosus*. These are virulence-associated proteins homologous to ORFs found on the F plasmid of *E. coli* (86). Furthermore, we found a KicA-like sequence; this protein is thought to suppress the killing function of the *kicB* gene product (49). The putative KicA-like protein in *P. luminescens* may, therefore, function as a toxin-antitoxin system for killing non-self bacteria, like the colicins and pycocins discussed above.

The sample sequence revealed five sequences that exhibit similarity to known two-component sensors: EnvZ, CheA, ExpA, BaeS, and TctE. Of these, only EnvZ and CheA have been characterized in detail. The *ompR-envZ* regulatory system has been shown to contribute to virulence in a number of enteric bacterial pathogens. For example, an isogenic *ompR* mutant of *Yersinia enterocolitica* exhibited increased sensitivity to high osmolarity, high temperature, and low pH and also offered partial protection against wild-type challenge in a murine yersiniosis model (43). The *ompR* and *envZ* signal transduction genes have also been cloned from another entomopathogenic nematode-associated bacterium, *Xenorhabdus nematophilus* (168). Deletion of *envZ* in a *Xenorhabdus* strain suggests that the gene regulates some outer membrane proteins during the stationary growth phase, implying that it has a potential role in virulence (see below). The CheA protein is required to initiate the response of the flagellar motor to the binding of stimulatory ligands to chemoreceptors during bacterial chemotaxis. The hit to ExpA is of great interest as this protein and its relatives appear to play a key role in regulating expression of a range of secreted virulence factors in different gram-negative bacteria. The relatives include SirA in *Salmo-*

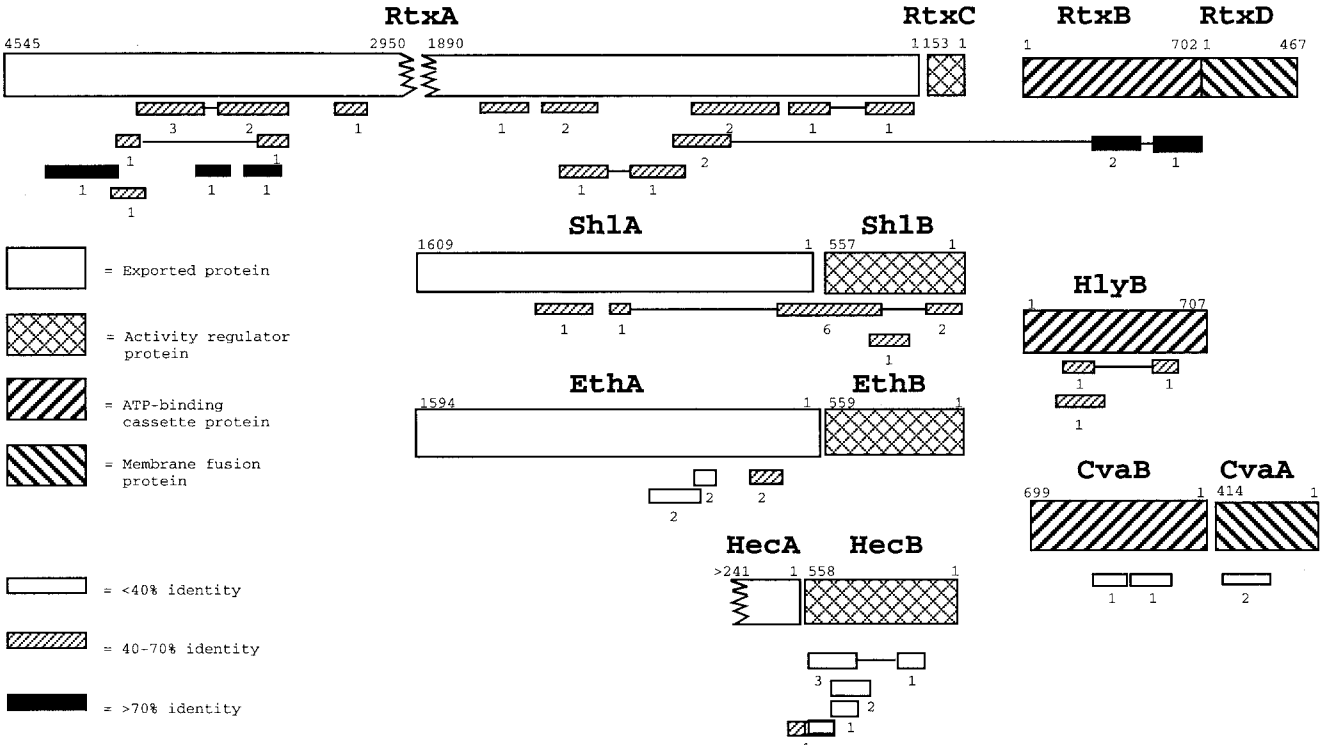

FIG. 3. Inferred genomic organization of different putative Rtx-like operons (*rtx*, *shl*, *eth*, *hec*, *hly*, and *cva*) from the sample sequence. The relative predicted positions of sequencing hits are shown below each predicted locus, and the range of percent identity values is shown. The putative operons are shaded in order to indicate their potential functions as either an exported protein, a activity regulator, an ATP-binding protein, or a membrane fusion protein.

nella typhimurium, ExpA in *Erwinia* spp., and GacA in *Pseudomonas* spp. For example, in *S. typhimurium* SirA is needed for expression of the type III secretion apparatus. Furthermore, upon sensing of a mammalian microenvironment, SirA phosphorylation initiates a cascade of transcription factor synthesis that leads not only to invasion gene transcription but also to Ssp secretion and bacterial epithelial invasion (78). Deletion of such a locus from a *Photorhabdus* strain would, therefore, allow us to test the hypothesis that a similar system is effective for sensing the insect hemocoel and subsequently initiating virulence-associated transcription.

Locomotion, attachment, and invasion. During its complex life cycle, a *Photorhabdus* strain not only needs to detect which environment it is in (e.g., nematode gut versus insect hemocoel) but presumably also needs to recognize specific surfaces for attachment and potentially invasion. Although we have little understanding of when and where *Photorhabdus* strains go during the insect infection process, we do know that their titers in the hemolymph change rapidly (52) and that during the infection process the insect midgut is specifically destroyed (18). Thus, we do not know if the bacteria replicate in the insect hemocytes or if they invade the gut directly. However, even in the absence of substantial information concerning the basic biology of these organisms, we can make inferences about the likely infection process based on the array of genes that they carry which are putatively involved in locomotion and tissue-specific attachment and/or invasion. Most notable in the latter case are two hits to the *attachment invasion locus* (*ail*) found in *Yersinia* spp. (Table 6). In *Y. enterocolitica* this locus is responsible for the ability of the pathogen to cross the epithelium of the gut on its way to replicate in the reticuloendothelial system. Again, although we have no direct evidence that a putative homolog plays a similar role in *Photorhabdus* strains, we can test whether *P. luminescens* W14 can invade the gut (presumably from the hemocoel, not the lumen) and, if it can, whether deletion of the *ail*-like locus interferes with this ability. With respect to attachment, we also found a sequence similar to intimins, which are proteins homologous to the invasins of *Yersinia* spp. and which play a role in attachment and effacing of the brush border membrane (54).

Another class of proteins involved in recognition of specific tissues is the class that includes the fimbriae and the associated adhesins. It has been hypothesized that in *X. nematophilus* fimbriae are involved in establishment of the specific association between the bacterium and the nematode gut (52). In the *Photorhabdus* sample sequence, we detected numerous matches with sequences encoding fimbrial type 1 subunits, fimbrial chaperones, and the outer membrane ushers associated with fimbrial export and assembly (Table 6). Although it is difficult to predict from these sequence matches the likely fimbrial composition of *P. luminescens* W14, we found both *mrpC*-like and *mrpD*-like loci, which encode the outer membrane usher and fimbrial chaperone from the Mrp (mannose-resistant, *Proteus*-like) fimbriae of *Proteus mirabilis*, respectively, and also FimD-like ushers and FimC-like chaperones from *E. coli*. The FimD sequence also exhibits similarity to S fimbrial adhesins, filamentous hemaglutinin A, and bovine colonization factor, implying that it may also play a role in virulence-associated adhesion. A second indication that there is another group of genes involved in a diverse array of functions that include fimbrial biogenesis, protein secretion, and DNA uptake (68) is the presence of sequences similar to those encoding a prepilin type of leader peptidase. Again, the significance of the pres-

TABLE 5. Putative virulence factors, genes expressed in infection, and two-component sensors

Gene function	Organism	Accession no.	No. of hits[a]	BLASTX *E* value(s)[b]	Clone(s)	Reference
Bioluminescence						
Luciferase, beta subunit	*P. luminescens*	C35411	1 (0)	3e-66	00415	[c]79
UbiB, NAD(P) H-flavin reductase	*P. luminescens*	P43129	3 (0)	1e-39 to 1e-71	00817, 02182, 02377	193
LumQ transcriptional regulator (linked to *lux* operon)	*Photobacterium leiognathi*	Q5187	1 (0)	2e-04	00131	111
Proteases, peptidases, and lipases						
Triaglycerol lipase 1	*P. luminescens*	P40601	4 (2)	4e-13 to 9e-86	00639, 00639f, 01676, 01676f	181
YfiP, lipase	*Bacillus subtilis*	D78508	2 (0)	2e-04, 2e-05	00339, 02380	188
Lipase (cold adapted)	*Pseudomonas* sp.	AF034088	1 (0)	3e-17	00082	33
PldB, lysophospholipase L2	*E. coli*	P07000	2 (0)	2e-20, 3e-49	02035, 02458	95
Lys-X cysteine protease	*P. gingivalis*	U83995	2 (0)	0.9 (66), 6e-04	01699, 02180	109
YegQ, putative collagenase	*E. coli*	P76403	2 (0)	4e-17, 7e-31	01523, 01829	19
OpdA, oligopeptidase A	*E. coli*	A43329	1 (0)	4e-50	01206	35
PrtB, oligopeptidase B	*E. coli*	P24555	3 (0)	3e-22 to 4e-47	00189, 01995, 01877	81
Protease IV	*E. coli*	P08395	2 (0)	1e-07, 1e-32	01602, 02400	74
Peptidase B	*E. coli*	P37095	1 (0)	4e-34	00116	189
Putative peptidase	*E. coli*	AE000321	1 (0)	4e-58	00083	19
Yae1, hypothetical	*E. coli*	P37764	1 (0)	2e-48	01938	19
Metalloprotease	*Chlamydia pneumoniae*	AE001618		6e-11		Unp.[c]
Matrix metalloprotease type 1	*Gallus gallus*	AF062392	1 (0)	7e-07	01471	Unp.
Major seed albumin	Pea	P08688	1 (0)	5e-04	01183	Unp.
Matrix metalloprotease type 2	*Limulus*	A40774		0.001 (second score)		113
Chitinases (*N*-acetyl-beta-glucosamidase)						
ChB, chitobiase	*S. marcescens*	Q54468	3 (0)	7e-34 to 1e-81	00174, 00222, 01723	171
Exochitinase chitinase C-like	S endosymbiont	Y11391	4 (2)	0.07 (61) to 9e-27	00596, 01515f, 02526, 02526f	Unp.
Hemagglutinins						
Putative secreted protein	*Neisseria meningitidis*	AF030941	3 (0)	2e-09 to 2e-34	00904, 01541, 01709	Unp.
FhaB, filamentous hemmaglutinin B	*B. pertussis*	P12255		0.04 (35) to 7e-05		148
Hemagglutinin neuramidase	Newcastle disease virus	M22110	1 (0)	4.1 (40)	01642	120
PalL, PA-I galactophilic lectin (galactophilic hemagglutinin)	*P. aeruginosa*	Q05097	1 (0)	3e-06	01333	8
Hemolysins (non-RTX)						
Ybex, hemolysin	*E. coli*	P77392	2 (0)	1e-48, 2e-78	00477, 01423	135
Hemolysin erythrocyte lysis protein 2	*Prevotella intermedia*	AF052516	1 (0)	1e-24	01236	14
ADP-ribosyltransferases and *B. thuringiensis*						
Cytotoxic necrotizing factor type 2	*E. coli* 711	A55260	1 (0)	3e-13	01649	137
ToxA, exotoxin A	*P. aeruginosa*	P11439	1 (0)	0.8 (44)	00003	110
ToxA, exotoxin A	*C. difficile*	A37052	1 (0)	0.05 (32)	02134	154
Halovibrin	*V. fisheri*	U38815	1 (0)	6e-19	00730	147
ExsA, exoenzyme S synthesis regulatory protein	*P. aeruginosa*	P26993	1 (0)	6e-50	01619	53
VirF, virulence regulon transcriptional regulator	*Y. enterocolitica*	P13225		2e-44	(Second match)	53
B. thuringiensis delta-endotoxin	*B. thuringiensis*	L07025	2 (0)	8.5 (26), 0.8 (30)	01891, 01973	101
Other virulence-associated factors						
VacB, RNase II	*H. influenzae*	P44907	1 (0)	5e-33	01226	51
VacB, RNase II	*E. coli*	P21499	2 (1)	1e-28, 9e-70	01226f, 02032	172
VapD, virulence-associated protein D	*H. influenzae*	C64069	1 (0)	2e-13	00582	51
VapZ, virulence-associated protein A′	*D. nodosus*	Q46561	2 (0)	0.2 (32), 8e-12	01633, 01947	86
KicA, killing factor	*E. coli*	S43912	1 (0)	2e-60	01813	49
MviM, virulence factor	*E. coli*	D90805	1 (0)	2e-45	02187	1
Two-component sensors						
EnvZ, osmolarity sensor	*Y. enterocolitica*	Y08950	1 (0)	2e-71	00849	43
CheA, chemotaxis protein	*S. typhimurium*	P09384	1 (0)	5e-37	01857	164
ExpA	*E. carotovora*	X95564	2 (0)	0.1 (25), 1e-39	00055, 02245	45
TctE	*S. typhi*	AF029846	1 (0)	2e-61	01665	Unp.
BaeS, sensory kinase	*E. coli*	P30847	1 (0)	3e-34	01887	128
Outer membrane proteins						
OmpF, porin	*S. marcescens*	033980	1 (0)	4e-43	01121	73
OmpW, outer membrane protein W	*E. coli*	P21364	1 (0)	4e-20	01233	125
OprF, outer membrane porin F	*Pseudomonas fluorescens*	AF117969	1 (0)	5e-06	02506	Unp.

[a] Number of sequence hits. The numbers in parentheses are the numbers of flipped sequences (indicated by the suffix f in the clone designations).
[b] The numbers in parentheses are percent amino acid identities.
[c] Unp., unpublished data.

TABLE 6. Genes encoding proteins important in attachment and locomotion including fimbriae, pili, and adhesins

Gene function	Organism	Accession no.	No. of hits[a]	BLASTX *E* value[b]	Clone(s)	Reference
Virulence-associated attachment						
Ail, attachment invasion	*Y. enterocolitica*	P16454	1 (0)	2e-15	00876	124
Ail, attachment invasion	*Y. pseudotuberculosis*	L49439	1 (0)	5e-13	02118	190
Int, intimin (invasin)	*E. coli*	AF043226	1 (0)	3.2 (40)	01861	54
TcpJ, toxin coregulated pilus leader peptidase	*V. cholerae*	P27717	1 (0)	1.3 (28)	01360	87
Fimbriae, chaperones, ushers, and adhesins						
AtfA, subunit of type 1 fimbria	*P. mirabilis*	Z78535	1 (0)	1.5 (27)	00198	117
YehA, type 1 fimbrial protein	*E. coli*	P33340	1 (0)	0.03 (30)	01345	19
FimA, type 1 fimbrial subunit	*S. marcescens*	P22595	1 (0)	0.2 (38)	00054	130
FimB, recombinase	*E. coli*	S533063	1 (0)	3e-15	01186	Unp.[c]
LpfB, chaperone (FimC-like)	*S. typhimurium*	P43661	1 (1)	1e-23	01029f	13
Putative chaperone (FimC-like)	*E. coli*	P77249	1 (0)	2e-11	01718	Unp.
Fimbrial chaperone (type 1)	*E. coli*	L77091	1 (0)	3e-10	00929	Unp.
YraI, fimbrial chaperone	*E. coli*	P42914	1 (0)	1.9 (31)	01426	Unp.
MrpD, fimbrial chaperone	*P. mirabilis*	Z32686	1 (0)	4e-77	02555	10
FimD, outer membrane usher	*S. typhimurium*	P37924	1 (1)	2e-16	00377f	Unp.
FimD, outer membrane usher	*E. coli*	P30130	2 (0)	3e-17, 4e-22	01205, 01626	92
HtrA, FimD-like usher	*E. coli*	P33129	1 (0)	1e-29	00506	145
MrpC, outer membrane usher	*P. mirabilis*	Z32686	2 (0)	9e-45, 9e-51	00479, 00155	10
Caf1A, F1 capsule anchoring (adhesin-like)	*Y. pestis*	P26949	1 (0)	7e-12	00377	83
S fimbrial adhesin	*E. coli*	1713397E	1 (0)	8e-24	01362	Unp.
Putative adhesin	*H. influenzae*	AF053125	1 (0)	2e-31	01905	114
B precursor (Zn binding)						
Fibronectin-binding protein B	*E. coli*	D90745	1 (0)	3e-13	01325	135
Prepilin peptidase and dependent protein						
TapD, prepilin peptidase type IV	*Aerumonas salmonicida*	AF059249	1 (0)	5e-27	02076	Unp.
Prepilin peptidase type IV	*Pseudomonas stutzeri*	AJ132364	1 (0)	0.004	01150	Unp.
PpdD, prepilin peptidase-dependent protein	*E. coli*	P36647	1 (0)	3e-09	01704	183
Flh and Fli						
Flagellar hook-associated protein 2 (FliD-like)	*X. nematophilus*	X91047	2 (0)	9e-55, 4e-63	01895, 01903	59
FliF, flagellar M-ring protein	*E. coli*	P25798	1 (0)	4e-68	00243	150
FliL, flagellar protein	*S. typhimurium*	P26417	1 (0)	2e-33	02318	89
FliQ, flagellar protein	*S. typhimurium*	P54701	1 (0)	2e-22	01686	Unp.
FliS, flagellar protein	*S. typhimurium*	P26609	1 (0)	1e-30	00187	88
FliZ, flagellar protein	*S. typhimurium*	AB010947	1 (0)	2e-28	02310	75
FlhA, flagellar biosynthesis	*P. mirabilis*	Q51910	1 (0)	3e-57	00007	Unp.
FlhE, flagellar protein	*E. coli*	P76297	1 (0)	0.19 (33)	01798	76
FlgA, flagellar basal body P-ring formation protein	*P. mirabilis*	U82214	1 (0)	3e-26	00676	64
FlgI, flagellar P-ring protein	*Agrobacterium*	Q44340	1 (0)	0.9 (37)	02471	38
FlgL, flagellar hook-associated protein (HAP3)	*S. typhimurium*	P13326	1 (0)	2e-08	01534	69
FlgN, flagellar synthesis protein	*P. mirabilis*	U82214	1 (0)	2e-32	00234	64
Flagellum-specific ATP synthase	*E. coli*	P52612	1 (0)	2e-68	01021	76

[a] Number of sequence hits. The numbers in parentheses are the numbers of flipped sequences (indicated by the suffix f in the clone designations).
[b] The numbers in parentheses are percent amino acid identities.
[c] Unp., unpublished data.

ence of these sequences in *Photorhabdus* sp. is not clear, but this topic warrants further investigation.

Flagella are important in bacterial locomotion, and phase I *Xenorhabdus* cells exhibit swarming motility when they are grown on suitable solid media (52). Correspondingly, extracts from phase I variants appear to contain flagellar filaments (flagellin), whereas phase II cells do not (52). Although the molecular mechanism of this defect in flagellin synthesis is unclear, we found several ORFs in both *fli*-like and *flh*-like operons in *P. luminescens* W14 (Table 6). These ORFs include the FliD-like hook-associated protein 2 previously cloned from *X. nematophilus* (59), which is differentially transcribed in the two phase variants. Previous experiments have shown that insertion of a transposon into the *flgN* gene of *P. mirabilis* resulted in a mutant which was still motile but had lost the ability to swarm (64). This suggests that specific flagella are independently responsible for the swarming and motility phenotypes. Identification of the genes encoding these two classes of flagella in *P. luminescens* may, therefore, enable us to elucidate not only what types of flagella are produced by the bacterium but in which phase variants they are expressed and what function they perform.

Finally, we found three different sequences that potentially encode outer membrane proteins (Omp). The outer membrane protein composition of *X. nematophilus* changes as the organism enters the stationary phase of growth, and the outer membrane proteins, which are thought to form pores, may be responsible for functions that are necessary for survival under stress conditions (52). For example, expression of cloned *ompF* of *S. marcescens* is increased in *E. coli* under high-osmolarity

TABLE 7. Secretion and transport, including Yop and low-calcium response-like sequences and ABC transporters

Gene function	Organism	Accession no.	No. of hits[a]	BLASTX *E* value(s)[b]	Clone(s)	Reference
Yops and low-calcium response-like stimulon (type III secretion)						
Yop37, outer membrane protein	*Y. enterocolitica* plasmid pYV	153665	1 (0)	5e-46	01174	11
YopT, Yop effector	*Y. enterocolitica*	AF12990	1 (0)	0.015 (31)	02522	Unp.
Invasin precursor/Yop1 adhesin	*Y. pseudotuberculosis*	P10858	1 (0)	4.1 (31)	02074	151
YscC, Yop secretion protein C	*Y. enterocolitica* plasmid pYV	Q01244	1 (0)	3e-23	02011	123
YscO, Yop secretion protein O	*Y. pestis*	AF020214	1 (0)	7e-13	01758	Unp.[c]
YscP, Yop secretion protein P	*Y. pestis*	P40295	1 (1)	1e-11	01758f	50
YscQ, Yop secretion protein Q	*Y. pseudotuberculosis*	P40296	1 (0)	1e-14	02311	16
SycN, YopN chaperone	*Y. enterocolitica* pYV03	M32097	1 (0)	4e-38	00164	178
General secretory pathway proteins						
General pathway protein F (cholera toxin secretion)	*V. cholerae*	P45780	1 (0)	2e-10	01002	Unp.
General pathway protein F	*Burkholderia pseudomallei*	AF110185	1 (0)	2.0 (27)	02509	40
ABC transporters: peptide and amino acid						
OppA, oligopeptide binding	*S. typhimurium*	P06202	2 (0)	7e-44, 6e-48	01442, 01703	66
OppB, oligopeptide transport	*E. coli*	P31132	1 (0)	1e-61	01507	84
DppA, dipeptide transport	*E. coli*	P23847	1 (0)	2e-19	00410	134
SapC, peptide transport	*E. coli*	Q47624	1 (0)	1e-68	01143	Unp.
ArtP, arginine transport	*E. coli*	P30858	1 (0)	1e-74	00618	19
ProU, glycine betaine transport	*E. coli*	P14175	1 (0)	2e-32	01868	163
GltL, amino acid transport	*H. influenzae*	P45022	1 (0)	3e-46	01437	51
TauB, taurine transport	*E. coli*	Q47538	1 (0)	8e-23	01789	176
Thiamine ABC transporter	*H. influenzae*	U32782	1 (0)	2e-32	00615	51
YvrO, amino acid transport	*Bacillus subtilis*	AJ223978	1 (0)	3e-08	00910	185
LivM, branched amino acids	*S. typhimurium*	P30296	1 (0)	3e-40	00622	118
CelB, cellobiose permease	*Borrelia burgdorferi*	AE000792	1 (0)	1e-25	02398	55
Sugar transport (Pts and Rbs)						
PtsG, glucose specific	*B. burgdorferi*	AE001166	1 (0)	1e-17	01900	Unp.
PTS, mannitol specific	*E. coli*	P00550	1 (0)	1e-35	01409	177
PTS, mannose subunit	*V. furnissii*	U65015	1 (0)	1e-35	01307	21
YidK, glucose transport-like	*E. coli*	P31448	3 (0)	4e-10 to 2e-38	01297, 01302, 01310	29
RbsA, ribose transport	*E. coli*	P04983	1 (0)	4e-89	01654	26
RbsC, ribose transport	*E. coli*	P04984	2 (0)	8e-11, 3e-28	00853, 00957	15
MalF, maltose transport	*Enterobacter aerogenes*	P18812	1 (0)	4e-38	01788	37
YbbA, heterocyst maturation	*H. influenzae*	P45247	1 (0)	4e-23	01616	51
YbbI, transcriptional regulator	*E. coli*	P77565	1 (0)	8e-27	02591	Unp.

[a] Number of sequence hits. The numbers in parentheses are the numbers of flipped sequences (indicated by the suffix f in the clone designations).
[b] The numbers in parentheses are percent amino acid identities.
[c] Unp., unpublished data.

conditions (73). It has also been hypothesized that the *X. nematophilus* outer membrane proteins play a role in specific interactions with the nematode host (52). Production of these proteins is regulated by EnvZ, as discussed above.

Secretion and transport. One of the most striking features of *P. luminescens* grown in a liquid culture is the large number of proteins that are secreted into the supernatant. Some of these proteins have been well characterized, including the Tc toxins, proteases, and lipases discussed above. However, most of the secreted proteins are poorly characterized, and, perhaps equally importantly, their mechanisms of export are not known. Thus, for example, the mechanism and timing of secretion of the Tc toxins in the insect host remain obscure. Below we discuss sequences similar to different types of secretion machinery, notably type III-like secretion systems and ABC transporters. We observed a series of hits to the Yop type III secretion system of *Yersinia* species, including sequences similar to both Yop proteins, Ysc secretion proteins, and Syc Yop-specific chaperones (Table 7). The *yop* virulon enables *Yersinia* cells to survive and multiply in the lymphoid tissues of their hosts (36). The Yop proteins are encoded on the pYV plasmid at the low-calcium-response locus, and virulent *Yersinia* cells secrete these virulence determinants when they are incubated at 37°C in the absence of Ca^{2+} ions. The Yop proteins themselves are involved in contact-dependent delivery of toxins and effector molecules. Thus, in *P. luminescens* they could potentially be responsible for delivering toxins to either the gut or the insect hemocytes. As discussed above, the *virF* virulence regulon transcriptional regulator (BLASTX *E* value, 2e-44) (Table 5) regulates production of Yop proteins. This gene is, therefore, a very interesting candidate for knockout in *P. luminescens*, as its loss may alter the pathogenesis of *Photorhabdus* cells with different insect tissues and potentially ascribe a function to the presence of the Yop-like sequences in strain W14.

In addition to contact-dependent secretion, the ABC transporters represent a large family of transporter systems with a diverse array of functions, including transport of peptides, amino acids, sugars, and metal ions. We, therefore, catalogued some of the sequences similar to ABC-like transporters (Table 7), and below we discuss some of their potential functions in *P. luminescens*. There were several sequences similar to peptide

TABLE 8. Polysaccharide biosynthesis, secretion, and recycling

Gene function	Organism	Accession no.	No. of hits[a]	BLASTX *E* value(s)	Clone(s)	Reference
Core LPS biosynthesis						
Putative glycosyltransferase	*S. marcescens*	U52844	2 (0)	1e-29, 7e-34	00798, 00944	61
Glycosyltransferase homolog	*B. pertussis*	S70676	1 (0)	1e-19	02392	5
TrsG, mannosyltransferase	*Y. enterocolitica*	S51266	1 (0)	3e-23, 3e-43	02267, 02222	158
RfaC, LPS heptosyltransferase 1	*E. coli*	P24173	1 (0)	8e-50	00086	31
RfaD, ADP-L-glycero-D-manno-heptose-6-epimerase	*H. influenzae*	P45048	1 (0)	2e-53	00230	51
RfbU-like, LPS biosynthesis	*Methanobacterium thermoautotrophicum*	AE000829	1 (0)	1e-16	00285	Unp.[b]
YfbE, spore coat polysaccharide	*E. coli*	P77690	2 (0)	7e-44, 2e-61	00452, 01453	Unp.
MulI, murein-lipoprotein	*Erwinia amylovora*	P02939	2 (0)	1e-30, 1e-30	01917, 02090	187
LpxA, UDP-*N*-acetylglucosamine acyltransferase	*P. mirabilis*	P72215	1 (0)	5e-8501015		Unp.
EnvA (LpxC), *N*-acetylglucosamine deacetylase	*E. coli*	P07652	1 (0)	3e-68	01963	94
RcsF, exopolysaccharide synthesis regulator	*E. coli*	P28633	1 (0)	2e-17	00997	57
Wza, polysaccharide export	*E. coli*	P76388	1 (0)	5e-81	00362	161
PulA, pullulanase	*Klebsiella aerogenes*	M16187	2 (0)	2e-18, 1e-71	00894, 00974	85
PulA, pullulanase	*Klebsiella pneumoniae*	P07206	1 (0)	5e-36	02084	98
GalE, UDP-glucose 4-epimerase	*Bacillus subtilis*	P55180	1 (0)	3e-05	02567	191
GalR, galactose operon repressor	*E. coli*	P03024	2 (0)	1e-09, 4e-23	01736, 01949	179
GalT, galactose-1-P uridyltransferase	*E. coli*	X06226	1 (0)	4e-28	02336	105

[a] Number of sequence hits. The numbers in parentheses are the numbers of flipped sequences (indicated by the suffix f in the clone designations).
[b] Unp., unpublished data.

and amino acid transporters. Two potential homologs, OppA and ProU, are of special interest. OppA is located in the periplasm and is required for uptake of peptide antibiotics in *E. coli* and *S. typhimurium* (66). ProU, the product of the *proU* locus (also found in both *E. coli* and *S. typhimurium*), is a high-affinity glycine betaine transport system which plays an important role in survival under osmotic stress conditions (163). There were also several sequences similar to various sugar transporters and their transcriptional regulators. Central among these was the bacterial phosphoenolpyruvate-dependent phosphotransferase system (PTS), which catalyzes cellular uptake and subsequent phosphorylation of carbohydrates and also plays a crucial role in the global regulation of various metabolic pathways (177). The presence of PTS-like sequences in *Photorhabdus* cells is potentially important because chitin-degrading bacteria, such as *Vibrio furnissii*, rely on PTSs in the chitin catabolic cascade (21), and *P. luminescens* may therefore utilize a similar system for degrading insect chitin.

Polysaccharide biosynthesis. Another striking feature of the *P. luminescens* culture supernatant is the large amount of LPS present. LPS production has been directly implicated in virulence in *P. luminescens*, as it has been in a wide range of other bacteria. For example, in *B. pertussis* the LPS is biologically active and is both toxic and immunogenic (5). LPS can also act as a recognition or binding site for extracellular agents. Thus, core LPS can act as a binding site for bacteriocins (alongside the outer membrane proteins OmpA and OmpF, as discussed above), while the *trsG* operon (Table 8) is required for biosynthesis of the bacteriophage Phi R1-37 receptor structures (158). The lipid A-core component of LPS is synthesized by sequential addition of sugars and fatty acids, and several sequences similar those involved in LPS biosynthesis were found in the sample sequence. These include *envA* (*lpxC*), which encodes an enzyme necessary for synthesis of the lipid A moiety (94), and *rfaC*, which is required for LPS inner-core synthesis (31). We also found genes likely to encode polysaccharide export functions, such as *rcsF*, which confers a mucoid phenotype (57), and *wza*, which encodes an outer membrane lipoprotein probably responsible for colanic acid (extracellular polysaccharide) export (161). Finally, genes encoding pullulanaselike proteins (starch-debranching enzymes) are also present; these proteins may play a role in recycling of the cell wall.

Iron acquisition and transport. As iron is often a rate-limiting growth factor in the host, many pathogenic bacteria have high-affinity iron-binding systems which can capture iron from host iron chelators. Thus, *P. luminescens* W14 has sequences which predict proteins similar to those involved in biosynthesis, transport, and reception of the siderophore yersiniabactin. Yersiniabactin (Ybt) has a high affinity for ferric iron, and similar siderophore-dependent iron transport systems are found in *Yersinia pestis*, *Yersinia pseudotuberculosis*, and *Y. enterocolitica* (140). A similar system may, therefore, also be used by *P. luminescens*. The *irp1* and *irp2* genes are required for yersiniabactin synthesis, as is *ybtE*, which encodes yersiniabactin dihydroxybenzoate ligase (Table 9). Transport of the iron-yersiniabactin complex back into the cell requires the TonB-dependent surface receptor FyuA, which may also be present in *P. luminescens* W14. This receptor is highly conserved and is found in all pesticin-sensitive bacteria, including *E. coli* (146). The sample sequence also contained hits to an R4-like ferric siderophore receptor from *E. coli*, which may perform a similar function in *P. luminescens*, and a putative operon (*pvcABCD*) involved in synthesis of the chromophore moiety of the *P. aeruginosa* siderophore pyoverdine (162).

P. luminescens, like *Yersinia* spp., also appears to contain alternative iron and hemin transport systems, as indicated by hits to genes similar to *yfeE*, the *yfeABCD* ferric iron uptake operon regulator, and members of the *hmu* hemin utilization system. The latter system is essential in *Y. pestis* for utilization

TABLE 9. Iron assimilation: ferric siderophore biosynthesis and transport and regulation of iron and other metals

Gene function	Organism	Accession no.	No. of hits[a]	BLASTX *E* value(s)	Clone(s)	Reference
Biosynthesis and reception of yersiniabactin-like siderophore						
Irp1, HMWP1-like	*Y. enterocolitica*	Y12527	12 (3)	4e-07 to 4e-35	00066, 00066f, 00245, 00269, 00498, 00647, 01200, 01491, 01491f, 01492, 01510, 01614f	139
Irp2, HMWP2-like	*Y. enterocolitica*	P48633	4 (2)	1e-15 to 5e-62	00062, 00062f, 00758, 00758f	139
Irp5, YbtE-like	*Y. pestis*	U50364	1 (0)	3e-59	01573	139
Ferric siderophore receptor-like						
FyuA, yersiniabactin receptor	*Y. enterocolitica*	P46360	1 (0)	2e-25	02508	146
R4, ferric siderophore receptor	*E. coli*	P27772	2 (0)	8e-06, 5e-30	00567, 01023	63
Pyoverdin siderophore synthesis						
PvcA, pyoverdin chromophore	*P. aeruginosa*	AF002222	1 (0)	1e-48	00220	162
PvcC, pyoverdin chromophore	*P. aeruginosa*	AF002222	1 (0)	1e-10	02521	162
Other iron and hemin transport systems						
YfeE, YfeABCD regulator	*Y. pestis*	Q56956	1 (0)	3e-32	01897	Unp.[b]
HmuR, outer membrane receptor	*Y. pestis*	Q56989	1 (0)	2e-62	002244	71
HmuS, transport protein	*Y. pestis*	Q56990	1 (0)	6e-57	00590	71
HemT, hemin binding periplasmic	*Y. enterocolitica*	X77867	1 (0)	8e-64	00516	165
HemT, hemin binding periplasmic	*Y. pestis*	Q56991		3e-45		71
FecA, outer membrane	*E. coli*	P13036	1 (0)	e-119	01194	159
FecC, cytosolic	*E. coli*	P15030	1 (0)	8e-38	02411	159
FecE, ATP binding	*Synechocystis* sp.	D90899	1 (0)	1e-12	00752	82
FeoB, ferrous iron transport	*E. coli*	P336650	2 (0)	5e-49, 3e-72	01462, 02346	80
Iron regulation and regulated proteins						
Fur, ferric uptake regulation	*E. coli*	P06975	1 (1)	2e-52	00174f	155
DtxR, diphtheria toxin repressor (iron dependent)	*Corynebacterium diphtheriae*	U20617	1 (0)	8e-18	02031	156
Hem biosynthesis						
Hem2, porphobilinogen synthase	*P. aeruginosa*	Q59643	1 (0)	7e-48	02146	Unp.
Hem6, coproporphyrinogen III oxidase, aerobic	*S. typhimurium*	P33771	2 (0)	8e-31, 2e-87	00810, 00883	186
HemE-like (DcuP) uroporphyrinogen decarboxylase	*E. coli*	P29680	1 (0)	2e-59	00202	131
HemN, coproporphyrinogen III oxidase, oxygen independent	*S. typhimurium*	P37129	1 (0)	8e-33	01830	186
HemN, coproporphyrinogonen III oxidase, oxygen independent	*Bacillus subtilis*	P54304	1 (0)	0.001	02315	70
HemY, protohem IX synthesis	*E. coli*	P09128	1 (0)	4e-14	00743	4
HemZ, ferrocheletase	*Y. enterocolitica*	P43413	1 (0)	8e-37	00814	165
NirJ-2, heme biosynthesis	*Archaeoglobus fulgidus*	AE000964	1 (0)	1e-05	00348	93
CysG, uroporphyrinogen III methylase, CG site 893	*E. coli*	P11098	2 (0)	1e-11, 5e-18	00955, 01812	138
CysI, sulfite reductase hemoprotein component	*E. coli*	M23008	1 (0)	3e-20	01919	136
Ferredoxin-like proteins						
Fer, ferredoxin	*E. coli*	P25528	1 (0)	2e-40	01022	167
YfhL, ferredoxin-like	*E. coli*	P52102	2 (0)	3e-33, 1e-38	01294, 01535	Unp.
HcaD-like, ferredoxin reductase	*Sphingomonas aromaticivorans*	AF079317	1 (0)	3e-08	01745	Unp.

[a] Number of sequence hits. The numbers in parentheses are the numbers of flipped sequences (indicated by the suffix f in the clone designations).
[b] Unp., unpublished data.

of free hemin and heme-protein complexes, which are the bacterium's sole sources of iron (71). *P. luminescens* also contains sequences similar to members of both the *feo* iron(II) (80) and *fec* iron(III) (159) transport systems. Finally, like numerous other gram-negative bacteria, *P. luminescens* also has a sequence similar to the *fur* gene sequence. This gene is involved in iron regulation, and in the presence of excess iron, the *fur* gene product generally represses expression of iron-regulated genes (140). Together, these sequences suggest that scavenging and transporting iron are important in *P. luminescens*, as they are in many pathogenic bacteria.

Extrachromosomal elements. Like the genomes of other bacteria, the *P. luminescens* sample sequence contains many sequences similar to sequences found in a wide range of phage and insertion sequence elements. These sequences are important because they may begin to explain how *P. luminescens*, an insect pathogen, acquired virulence factors previously associated only with vertebrate pathogenicity, such as sequences similar to the low-calcium-response stimulon from *Yersinia* discussed above. Numerous hits to tail proteins from P2-like bacteriophages (46) (P2, P4, 186, and HP1) and a range of other phage-related proteins were observed ($n = 51$). There were

TABLE 10. Extrachromosomal elements or inserted elements, including transposons and insertion sequence elements[a]

Gene function	Organism	Accession no.	No. of hits[b]	BLASTX *E* value(s)[c]	Clone(s)	Reference
Plasmid associated and plasmid stability						
Orf4, hypothetical	*Enterobacteriaceae* plasmid R100	S28661	1 (0)	4.4 (33)	01224	7
YY08, predicted Orf	*Methanococcus jannaschii* plasmid pURB801	Q60307	1 (0)	2.0 (26)	00214	27
ParD, stabilization factor	*E. coli* pRP4	P22995	1 (0)	2e-32	00808	149
StbB, stability protein	*P. syringae*	Q52562	1 (0)	1e-17	01196	Unp.[d]
StbD, stability protein	*Morganella morganii*	AF072126	1 (0)	7e-12	01412	65
Plasmid stability-like	*Thiobacillus ferrooxidans*	U73041	1 (0)	2.0 (50)	01275	42
Dma7, DNA adenine methylase	*E. coli* retron EC67	P21311	1 (0)	0.013, 8e-23	00173	72
Orf1, hypothetical	*E. coli* retron EC67	P21323	1 (0)	8e-23	01168	72
Replication, repair, transformation, and conjugation						
PriA, replication factor N′	*E. coli*	A35505	2 (0)	5e-69, 1e-87	01203, 01446	132
PriC, replication factor N″	*E. coli*	P23862	1 (0)	3e-07	01634	192
UvrA, ABC excinuclease A	*S. typhimurium*	P37434	1 (0)	4e-58	01967	Unp.
UvrB, ABC exinuclease B	*E. coli*	P07025	1 (0)	1e-95	02230	9
UvrC, ABC exinuclease C	*E. coli*	P07028	1 (0)	4e-28	01063	153
TfoX, DNA transformation	*H. influenzae*	P43779	1 (0)	3e-05	00715	194
ComE, DNA transformation	*H. influenzae*	P31772	1 (0)	2e-24	00510	102
DNA transformation-like	*H. influenzae*	JH0436	1 (0)	3e-26	01234	173
Ex5A, exodeoxyribonuclease	*E. coli*	P04993	3 (0)	7e-32 to 2e-70	01529, 01652, 01842	19
Restriction enzymes and their control						
Tlr1, type I restriction enzyme *Eco*R124II	*E. coli*	P10486	1 (0)	2e-84	02345	144
Tls1, type 1 restriction enzyme *Eco*RI24II specificity	*E. coli*	P10485	1 (0)	6e-30	00989	144
*Ngo*MI, restriction enzyme type II	*Neisseria gonorrhoeae*	P31032	1 (1)	3e-07	00484f	160
*Nae*I modification methylase cytosine specific	*Nocardia aerocolonigenes*	P50188	1 (1)	3e-20	01699f	169
*Bam*HI control element	*Bacillus amyloliquefaciens*	X55285	1 (0)	4e-06	00175	25

[a] Data for phage and phage-related proteins are not included.
[b] Number of sequence hits. The numbers in parentheses are the numbers of flipped sequences (indicated by the suffix f in the clone designations).
[c] The numbers in parentheses are percent amino acid identities.
[d] Unp., unpublished data.

also 10 hits (BLASTX *P* values, 0.01 to 2e-86) to products of the *integrase* (*int*) gene, which controls phage site-specific integration. Notably, in the range of phage homologies there were hits (although with relatively low significance [BLASTX *E* values, 0.3 to 5e-15]) to three different ORFs (ORFs 16, 20, and 25) of the *P. aeruginosa* cytotoxin converting phage Phi CTX (cholera toxin). We note that the *rtx* gene cluster is physically linked to the CTX toxin element in the *V. cholerae* genome (112). Therefore, it will be interesting to investigate whether this element is linked with the *rtxA*-like sequences found in *P. luminescens* W14, suggesting that it could have been responsible for horizontal transfer of the toxin-encoding genes. Numerous transposon-like sequences were also found ($n = 33$), including 10 hits to a transposase from plasmid ColIb-P9 (BLASTX *E* values, 1e-04 to 4e-86). Again, although these sequences indicate that transfer events occurred, it is not known how long these transposons have been present and if any of them have retained functionality. Finally, the *P. luminescens* W14 genome contains numerous sequences related to sequences involved in plasmid maintenance and stability (Table 10). However, we cannot at this stage distinguish which of these sequences are plasmid encoded (plasmids have been found previously in *Xenorhabdus* spp. [103]) and which are chromosomal. The presence of these sequences, therefore, raises the possibility of plasmid maintenance in *P. luminescens* W14 but is not strictly indicative.

Conclusions. *P. luminescens* has a life cycle which introduces it into a diverse array of environments, and in only one of these environments, the insect environment, is the bacterium pathogenic. The sample sequence of strain W14 revealed sequences similar to the sequences of a diverse array of potential virulence factor-encoding genes, including the genes for several classes of toxins, proteases, lipases, and LPS. It also gave us some indication of the diversity of the transport and metabolic systems present. Furthermore, *Photorhabdus* spp. also seem to share potential virulence factors (Yops, a yersiniabactin-like siderophore, and the low-calcium-response stimulon) with distantly related vertebrate pathogens, such as members of the genus *Yersinia*. This hypothesis is supported by the presence of numerous phagelike and transposon-like sequences in the *P. luminescens* genome. The potential for horizontal transfer raises the intriguing possibility that the virulence factors present in invertebrate pathogens may also be present in vertebrate pathogens. Given the far greater diversity of invertebrates and, potentially, their associated pathogens, this raises interesting questions about the diversity and origins of potential vertebrate virulence factors. In relation to *P. luminescens* itself, complete elucidation of the genome sequence of strain

W14 and other strains should allow us to begin to understand the roles of individual genes via targeted disruption and to begin to compare the diversity of virulence factors found in different invertebrate pathogens. Our findings are consistent with the hypothesis of Burland et al. (30), who hypothesized that all of the pathogenic genes shared by enteric bacteria form a pool or "pathosphere"; however, here we emphasize that the pool must be extended to include both invertebrate and vertebrate pathogens. Furthermore, as invertebrates evolved before vertebrates, this also raises the interesting possibility that pathogens such as *P. luminescens* include the progenitors of virulence factors in vertebrate pathogens.

ACKNOWLEDGMENTS

This work was supported by grants to R.F.C from the BBSRC, the Wellcome Trust (JIF), and DowAgroSciences, which we thank for their interest and support.

We thank all of the technical staff of the Wisconsin Genome Project for their help with sequencing and assembly. We thank David Clarke for comments on the manuscript.

REFERENCES

1. **Aiba, H., T. Baba, K. Hayashi, T. Inada, K. Isono, T. Itoh, H. Kasai, K. Kashimoto, S. Kimura, M. Kitakawa, M. Kitagawa, K. Makino, T. Miki, K. Mizobuchi, H. Mori, T. Mori, K. Motomura, S. Nakade, Y. Nakamura, H. Nashimoto, Y. Nishio, T. Oshima, N. Saito, G. Sampei, T. Horiuchi, et al.** 1996. A 570-kb DNA sequence of the *Escherichia coli* K-12 genome corresponding to the 28.0-40.1 min region on the linkage map. DNA Res. **3:**363–377.
2. **Akatsuka, H., E. Kawai, K. Omori, and T. Shibatani.** 1995. The three genes *lipB*, *lipC*, and *lipD* involved in the extracellular secretion of the *Serratia marcescens* lipase which lacks an N-terminal signal peptide. J. Bacteriol. **177:**6381–6389.
3. **Albertini, A. M., T. Caramori, F. Scoffone, C. Scotti, and A. Galizzi.** 1995. Sequence around the 159 degree region of the *Bacillus subtilis* genome: the *pksX* locus spans 33.6 kb. Microbiology **141:**299–309.
4. **Alefounder, P. R., C. Abell, and A. R. Battersby.** 1988. The sequence of *hemC*, *hemD* and two additional *E. coli* genes. Nucleic Acids Res. **16:**9871.
5. **Allen, A., and D. Maskell.** 1996. The identification, cloning and mutagenesis of a genetic locus required for lipopolysaccharide biosynthesis in *Bordetella pertussis*. Mol. Microbiol. **19:**37–52.
6. **Aono, R., T. Negishi, and H. Nakajima.** 1994. Cloning of organic solvent tolerance gene *ostA* that determines *n*-hexane tolerance level in *Escherichia coli*. Appl. Environ. Microbiol. **60:**4624–4626.
7. **Armstrong, K. A., H. Ohtsubo, W. R. Bauer, Y. Yoshioka, C. Miyazaki, Y. Maeda, and E. Ohtsubo.** 1986. Characterization of the gene products produced in minicells by pSM1, a derivative of R100. Mol. Gen. Genet. **205:** 56–65.
8. **Avichezer, D., N. Gilboa-Garber, N. C. Garber, and D. J. Katcoff.** 1994. *Pseudomonas aeruginosa* PA-I lectin gene molecular analysis and expression in *Escherichia coli*. Biochim. Biophys. Acta **1218:**11–20.
9. **Backendorf, C., H. Spaink, A. P. Barbeiro, and P. van de Putte.** 1986. Structure of the *uvrB* gene of *Escherichia coli*. Homology with other DNA repair enzymes and characterization of the *uvrB5* mutation. Nucleic Acids Res. **14:**2877–2890.
10. **Bahrani, F. K., and H. L. Mobley.** 1994. *Proteus mirabilis* MR/P fimbrial operon: genetic organization, nucleotide sequence, and conditions for expression. J. Bacteriol. **176:**3412–3419.
11. **Barteneva, N. S., A. G. Evstafieva, V. N. Gorelov, and B. E. Wenzel.** 1994. Identification and sequencing of a plasmid (pYV96)-encoded gene product of *Yersinia enterocolitica* recognized by antibodies in sera of patients with autoimmune thyroid disease. Ann. N. Y. Acad. Sci. **730:**345–347.
12. **Bauer, D. W., Z. M. Wei, S. V. Beer, and A. Collmer.** 1995. *Erwinia chrysanthemi* harpinEch: an elicitor of the hypersensitive response that contributes to soft-rot pathogenesis. Mol. Plant-Microbe Interact. **8:**484–491.
13. **Baumler, A. J., and F. Heffron.** 1995. Identification and sequence analysis of *lpfABCDE*, a putative fimbrial operon of *Salmonella typhimurium*. J. Bacteriol. **177:**2087–2097.
14. **Beem, J. E., W. E. Nesbitt, and K. P. Leung.** 1999. Cloning of *Prevotella intermedia* loci demonstrating multiple hemolytic domains. Oral Microbiol. Immunol. **14:**143–152.
15. **Bell, A. W., S. D. Buckel, J. M. Groarke, J. N. Hope, D. H. Kingsley, and M. A. Hermodson.** 1986. The nucleotide sequences of the *rbsD*, *rbsA*, and *rbsC* genes of *Escherichia coli* K12. J. Biol. Chem. **261:**7652–7658.
16. **Bergman, T., K. Erickson, E. Galyov, C. Persson, and H. Wolf-Watz.** 1994. The *lcrB* (*yscN*/*U*) gene cluster of *Yersinia pseudotuberculosis* is involved in Yop secretion and shows high homology to the *spa* gene clusters of *Shigella flexneri* and *Salmonella typhimurium*. J. Bacteriol. **176:**2619–2626.
17. **Bernhard, F., G. Demel, K. Soltani, H. V. Dohren, and V. Blinov.** 1996. Identification of genes encoding for peptide synthetases in the gram-negative bacterium *Lysobacter* sp. ATCC 53042 and the fungus *Cylindrotrichum oligospermum*. DNA Seq. **6:**319–330.
18. **Blackburn, M., E. Golubeva, D. Bowen, and R. H. ffrench-Constant.** 1998. A novel insecticidal toxin from *Photorhabdus luminescens*, Toxin complex a (Tca), and its histopathological effects on the midgut of *Manduca sexta*. Appl. Environ. Microbiol. **64:**3036–3041.
19. **Blattner, F. R., G. Plunkett 3rd, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao.** 1997. The complete genome sequence of *Escherichia coli* K-12. Science **277:**1453–1474.
20. **Borchert, S., S. S. Patil, and M. A. Marahiel.** 1992. Identification of putative multifunctional peptide synthetase genes using highly conserved oligonucleotide sequences derived from known synthetases. FEMS Microbiol. Lett. **71:**175–180.
21. **Bouma, C. L., and S. Roseman.** 1996. Sugar transport by the marine chitinolytic bacterium *Vibrio furnissii*. Molecular cloning and analysis of the mannose/glucose permease. J. Biol. Chem. **271:**33468–33475.
22. **Bowen, D., T. A. Rocheleau, M. Blackburn, O. Andreev, E. Golubeva, R. Bhartia, and R. H. ffrench-Constant.** 1998. Insecticidal toxins from the bacterium *Photorhabdus luminescens*. Science **280:**2129–2132.
23. **Bowen, D. J., and J. C. Ensign.** 1998. Purification and characterization of a high-molecular-weight insecticidal protein complex produced by the entomopathogenic bacterium *Photorhabdus luminescens*. Appl. Environ. Microbiol. **64:**3029–3035.
24. **Boyd, C., and N. T. Keen.** 1993. Characterization of the *prtA* and *prtB* genes of *Erwinia chrysanthemi* EC16. Gene **133:**115–118.
25. **Brooks, J. E., P. D. Nathan, D. Landry, L. A. Sznyter, P. Waite-Rees, C. L. Ives, L. S. Moran, B. E. Slatko, and J. S. Benner.** 1991. Characterization of the cloned *Bam*HI restriction modification system: its nucleotide sequence, properties of the methylase, and expression in heterologous hosts. Nucleic Acids Res. **19:**841–850.
26. **Buckel, S. D., A. W. Bell, J. K. Rao, and M. A. Hermodson.** 1986. An analysis of the structure of the product of the *rbsA* gene of *Escherichia coli* K12. J. Biol. Chem. **261:**7659–7662.
27. **Bult, C. J., O. White, G. J. Olsen, L. Zhou, R. D. Fleischmann, G. G. Sutton, J. A. Blake, L. M. FitzGerald, R. A. Clayton, J. D. Gocayne, A. R. Kerlavage, B. A. Dougherty, J. F. Tomb, M. D. Adams, C. I. Reich, R. Overbeek, E. F. Kirkness, K. G. Weinstock, J. M. Merrick, A. Glodek, J. L. Scott, N. S. M. Geoghagen, and J. C. Venter.** 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science **273:**1058–1073.
28. **Burland, V., D. L. Daniels, G. D. Plunkett, and F. R. Blattner.** 1993. Genome sequencing on both strands: the Janus strategy. Nucleic Acids Res. **21:**3385–3390.
29. **Burland, V., G. D. Plunkett, D. L. Daniels, and F. R. Blattner.** 1993. DNA sequence and analysis of 136 kilobases of the *Escherichia coli* genome: organizational symmetry around the origin of replication. Genomics **16:** 551–561.
30. **Burland, V., Y. Shao, N. T. Perna, G. Plunkett, H. J. Sofia, and F. R. Blattner.** 1998. The complete DNA sequence and analysis of the large virulence plasmid of *Escherichia coli* O157:H7. Nucleic Acids Res. **26:**4196–4204.
31. **Chen, L., and W. G. Coleman, Jr.** 1993. Cloning and characterization of the *Escherichia coli* K-12 *rfa-2* (*rfaC*) gene, a gene required for lipopolysaccharide inner core synthesis. J. Bacteriol. **175:**2534–2540.
32. **Chen, R., and U. Henning.** 1987. Nucleotide sequence of the gene for the peptidoglycan-associated lipoprotein of *Escherichia coli* K12. Eur. J. Biochem. **163:**73–77.
33. **Choo, D. W., T. Kurihara, T. Suzuki, K. Soda, and N. Esaki.** 1998. A cold-adapted lipase of an Alaskan psychrotroph, *Pseudomonas* sp. strain B11-1: gene cloning and enzyme purification and characterization. Appl. Environ. Microbiol. **64:**486–491.
34. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry 3rd, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, B. G. Barrell, et al.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393:**537–544.
35. **Conlin, C. A., N. J. Trun, T. J. Silhavy, and C. G. Miller.** 1992. *Escherichia coli prlC* encodes an endopeptidase and is homologous to the *Salmonella typhimurium opdA* gene. J. Bacteriol. **174:**5881–5887.
36. **Cornelis, G. R., and H. Wolf-Watz.** 1997. The *Yersinia* Yop virulon: a bacterial system for subverting eukaryotic cells. Mol. Microbiol. **23:**861–867.
37. **Dahl, M. K., E. Francoz, W. Saurin, W. Boos, M. D. Manson, and M. Hofnung.** 1989. Comparison of sequences from the *malB* regions of *Salmonella typhimurium* and *Enterobacter aerogenes* with *Escherichia coli* K12: a potential new regulatory site in the interoperonic region. Mol. Gen. Genet. **218:**199–207.

38. **Deakin, W. J., C. S. Furniss, V. E. Parker, and C. H. Shaw.** 1997. Isolation and characterisation of a linked cluster of genes from *Agrobacterium tumefaciens* encoding proteins involved in flagellar basal-body structure. Gene **189:**135–137.
39. **de Crecy-Lagard, V., V. Blanc, P. Gil, L. Naudin, S. Lorenzon, A. Famechon, N. Bamas-Jacques, J. Crouzet, and D. Thibaut.** 1997. Pristinamycin I biosynthesis in *Streptomyces pristinaespiralis*: molecular characterization of the first two structural peptide synthetase genes. J. Bacteriol. **179:**705–713.
40. **DeShazer, D., P. J. Brett, M. N. Burtnick, and D. E. Woods.** 1999. Molecular characterization of genetic loci required for secretion of exoproducts in *Burkholderia pseudomallei*. J. Bacteriol. **181:**4661–4664.
41. **Dittmann, E., B. A. Neilan, M. Erhard, H. von Dohren, and T. Borner.** 1997. Insertional mutagenesis of a peptide synthetase gene that is responsible for hepatotoxin production in the cyanobacterium *Microcystis aeruginosa* PCC 7806. Mol. Microbiol. **26:**779–787.
42. **Dominy, C. N., S. M. Deane, and D. E. Rawlings.** 1997. A geographically widespread plasmid from *Thiobacillus ferrooxidans* has genes for ferredoxin-, FNR-, prismane- and NADH-oxidoreductase-like proteins which are also located on the chromosome. Microbiology **143:**3123–3136.
43. **Dorrell, N., S. R. Li, P. H. Everest, G. Dougan, and B. W. Wren.** 1998. Construction and characterization of a *Yersinia enterocolitica* O:8 *ompR* mutant. FEMS Microbiol. Lett. **165:**145–151.
44. **Duport, C., C. Baysse, and Y. Michel-Briand.** 1995. Molecular characterization of pyocin S3, a novel S-type pyocin from *Pseudomonas aeruginosa*. J. Biol. Chem. **270:**8920–8927.
45. **Eriksson, A. R., R. A. Andersson, M. Pirhonen, and E. T. Palva.** 1998. Two-component regulators involved in the global control of virulence in *Erwinia carotovora* subsp. *carotovora*. Mol. Plant-Microbe Interact. **11:**743–752.
46. **Esposito, D., W. P. Fitzmaurice, R. C. Benjamin, S. D. Goodman, A. S. Waldman, and J. J. Scocca.** 1996. The complete nucleotide sequence of bacteriophage HP1 DNA. Nucleic Acids Res. **24:**2360–2368.
47. **Farris, M., A. Grant, T. B. Richardson, and C. D. O'Connor.** 1998. BipA: a tyrosine-phosphorylated GTPase that mediates interactions between enteropathogenic *Escherichia coli* (EPEC) and epithelial cells. Mol. Microbiol. **28:**265–279.
48. **Felmlee, T., S. Pellett, and R. A. Welch.** 1985. Nucleotide sequence of an *Escherichia coli* chromosomal hemolysin. J. Bacteriol. **163:**94–105.
49. **Feng, J., K. Yamanaka, H. Niki, T. Ogura, and S. Hiraga.** 1994. New killing system controlled by two genes located immediately upstream of the *mukB* gene in *Escherichia coli*. Mol. Gen. Genet. **243:**136–147.
50. **Fields, K. A., G. V. Plano, and S. C. Straley.** 1994. A low-Ca^{2+} response (LCR) secretion (*ysc*) locus lies within the *lcrB* region of the LCR plasmid in *Yersinia pestis*. J. Bacteriol. **176:**569–579.
51. **Fleishmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J. F. Tomb, B. A. Dougherty, J. M. Merrick, et al.** 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269:**496–512.
52. **Forst, S., and K. Nealson.** 1996. Molecular biology of the symbiotic-pathogenic bacteria *Xenorhabdus* spp. and *Photorhabdus* spp. Microbiol. Rev. **60:**21–43.
53. **Frank, D. W., and B. H. Iglewski.** 1991. Cloning and sequence analysis of a trans-regulatory locus required for exoenzyme S synthesis in *Pseudomonas aeruginosa*. J. Bacteriol. **173:**6460–6468.
54. **Frankel, G., D. C. Candy, P. Everest, and G. Dougan.** 1994. Characterization of the C-terminal domains of intimin-like proteins of enteropathogenic and enterohemorrhagic *Escherichia coli*, *Citrobacter freundii*, and *Hafnia alvei*. Infect. Immun. **62:**1835–1842.
55. **Fraser, C. M., S. Casjens, W. M. Huang, G. G. Sutton, R. Clayton, R. Lathigra, O. White, K. A. Ketchum, R. Dodson, E. K. Hickey, M. Gwinn, B. Dougherty, J. F. Tomb, R. D. Fleishmann, D. Richardson, J. Peterson, A. R. Kerlavage, J. Quackenbush, S. Salzberg, M. Hanson, R. van Vugt, N. Palmer, M. D. Adams, J. Gocayne, J. C. Venter, et al.** 1997. Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. Nature **390:**580–586.
56. **Garcia del Portillo, F., M. A. de Pedro, and J. A. Ayala.** 1991. Identification of a new mutation in *Escherichia coli* that suppresses a *pbpB* (Ts) phenotype in the presence of penicillin-binding protein 1B. FEMS Microbiol. Lett. **68:**7–13.
57. **Gervais, F. G., and G. R. Drapeau.** 1992. Identification, cloning, and characterization of *rcsF*, a new regulator gene for exopolysaccharide synthesis that suppresses the division mutation ftsZ84 in *Escherichia coli* K-12. J. Bacteriol. **174:**8016–8022.
58. **Gilson, L., H. K. Mahanty, and R. Kolter.** 1990. Genetic analysis of an MDR-like export system: the secretion of colicin V. EMBO J. **9:**3875–3894.
59. **Givaudan, A., A. Lanois, and N. Boemare.** 1996. Cloning and nucleotide sequence of a flagellin encoding genetic locus from *Xenorhabdus nematophilus*: phase variation leads to differential transcription of two flagellar genes (*fliCD*). Gene **183:**243–253.
60. **Goncharoff, P., S. Saadi, C. H. Chang, L. H. Saltman, and D. H. Figurski.** 1991. Structural, molecular, and genetic analysis of the *kilA* operon of broad- host-range plasmid RK2. J. Bacteriol. **173:**3463–3477.
61. **Guasch, J. F., N. Pique, N. Climent, S. Ferrer, S. Merino, X. Rubires, J. M. Tomas, and M. Regue.** 1996. Cloning and characterization of two *Serratia marcescens* genes involved in core lipopolysaccharide biosynthesis. J. Bacteriol. **178:**5741–5747.
62. **Guenzi, E., G. Galli, I. Grgurina, D. C. Gross, and G. Grandi.** 1998. Characterization of the syringomycin synthetase gene cluster: a link between prokaryotic and eukaryotic peptide synthetases. J. Biol. Chem. **49:**32857–32863.
63. **Guyer, D. M., J. S. Kao, and H. L. Mobley.** 1998. Genomic analysis of a pathogenicity island in uropathogenic *Escherichia coli* CFT073: distribution of homologous sequences among isolates from patients with pyelonephritis, cystitis, and catheter-associated bacteriuria and from fecal samples. Infect. Immun. **66:**4411–4417.
64. **Gygi, D., G. Fraser, A. Dufour, and C. Hughes.** 1997. A motile but nonswarming mutant of *Proteus mirabilis* lacks FlgN, a facilitator of flagella filament assembly. Mol. Microbiol. **25:**597–604.
65. **Hayes, F.** 1998. A family of stability determinants in pathogenic bacteria. J. Bacteriol. **180:**6415–6418.
66. **Hiles, I. D., and C. F. Higgins.** 1986. Peptide uptake by *Salmonella typhimurium*. The periplasmic oligopeptide-binding protein. Eur. J. Biochem. **158:**561–567.
67. **Hirono, I., N. Tange, and T. Aoki.** 1997. Iron-regulated haemolysin gene from *Edwardsiella tarda*. Mol. Microbiol. **24:**851–856.
68. **Hobbs, M., and J. S. Mattick.** 1993. Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentous phage and protein-secretion apparatus: a general system for the formation of surface-associated protein complexes. Mol. Microbiol. **10:**233–243.
69. **Homma, M., D. J. DeRosier, and R. M. Macnab.** 1990. Flagellar hook and hook-associated proteins of *Salmonella typhimurium* and their relationship to other axial components of the flagellum. J. Mol. Biol. **213:**819–832.
70. **Homuth, G., M. Heinemann, U. Zuber, and W. Schumann.** 1996. The genes of *lepA* and *hemN* form a bicistronic operon in *Bacillus subtilis*. Microbiology **142:**1641–1649.
71. **Hornung, J. M., H. A. Jones, and R. D. Perry.** 1996. The *hmu* locus of *Yersinia pestis* is essential for utilization of free haemin and haem-protein complexes as iron sources. Mol. Microbiol. **20:**725–739.
72. **Hsu, M. Y., M. Inouye, and S. Inouye.** 1990. Retron for the 67-base multicopy single-stranded DNA from *Escherichia coli*: a potential transposable element encoding both reverse transcriptase and Dam methylase functions. Proc. Natl. Acad. Sci. USA **87:**9454–9458.
73. **Hutsul, J. A., and E. Worobec.** 1997. Molecular characterization of the *Serratia marcescens* OmpF porin, and analysis of *S. marcescens* OmpF and OmpC osmoregulation. Microbiology **143:**2797–2806.
74. **Ichihara, S., T. Suzuki, M. Suzuki, and S. Mizushima.** 1986. Molecular cloning and sequencing of the *sppA* gene and characterization of the encoded protease IV, a signal peptide peptidase, of *Escherichia coli*. J. Biol. Chem. **261:**9405–9411.
75. **Ikebe, T., S. Iyoda, and K. Kutsukake.** 1999. Structure and expression of the *fliA* operon of *Salmonella typhimurium*. Microbiology **145:**1389–1396.
76. **Itoh, T., H. Aiba, T. Baba, K. Hayashi, T. Inada, K. Isono, H. Kasai, S. Kimura, M. Kitakawa, M. Kitagawa, K. Makino, T. Miki, K. Mizobuchi, H. Mori, T. Mori, K. Motomura, S. Nakade, Y. Nakamura, H. Nashimoto, Y. Nishio, T. Oshima, N. Saito, G. Sampei, Y. Seki, T. Horiuchi, et al.** 1996. A 460-kb DNA sequence of the *Escherichia coli* K-12 genome corresponding to the 40.1-50.0 min region on the linkage map. DNA Res. **3:**379–392.
77. **Jackowski, S., P. D. Jackson, and C. O. Rock.** 1994. Sequence and function of the *aas* gene in *Escherichia coli*. J. Biol. Chem. **269:**2921–2928.
78. **Johnston, C., D. A. Pegues, C. J. Hueck, A. Lee, and S. I. Miller.** 1996. Transcriptional activation of *Salmonella typhimurium* invasion genes by a member of the phosphorylated response-regulator superfamily. Mol. Microbiol. **22:**715–727.
79. **Johnston, T. C., E. B. Rucker, L. Cochrum, K. S. Hruska, and V. Vandegrift.** 1990. The nucleotide sequence of the *luxA* and *luxB* genes of *Xenorhabdus luminescens* HM and a comparison of the amino acid sequences of luciferases from four species of bioluminescent bacteria. Biochem. Biophys. Res. Commun. **170:**407–415.
80. **Kammler, M., C. Schon, and K. Hantke.** 1993. Characterization of the ferrous iron uptake system of *Escherichia coli*. J. Bacteriol. **175:**6212–6219.
81. **Kanatani, A., T. Masuda, T. Shimoda, F. Misoka, X. S. Lin, T. Yoshimoto, and D. Tsuru.** 1991. Protease II from *Escherichia coli*: sequencing and expression of the enzyme gene and characterization of the expressed enzyme. J. Biochem. (Tokyo) **110:**315–320.
82. **Kaneko, T., S. Sato, H. Kotani, A. Tanaka, E. Asamizu, Y. Nakamura, N. Miyajima, M. Hirosawa, M. Sugiura, S. Sasamoto, T. Kimura, T. Hosouchi, A. Matsuno, A. Muraki, N. Nakazaki, K. Naruo, S. Okumura, S. Shimpo, C. Takeuchi, T. Wada, A. Watanabe, M. Yamada, M. Yasuda, and S. Tabata.** 1996. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res. **3:**109–136.
83. **Karlyshev, A. V., E. E. Galyov, O. Smirnov, A. P. Guzayev, V. M. Abramov, and V. P. Zav'yalov.** 1992. A new gene of the *f1* operon of *Y. pestis* involved

in the capsule biogenesis. FEBS Lett. **297:**77–80.
84. **Kashiwagi, K., Y. Yamaguchi, Y. Sakai, H. Kobayashi, and K. Igarashi.** 1990. Identification of the polyamine-induced protein as a periplasmic oligopeptide binding protein. J. Biol. Chem. **265:**8387–8391.
85. **Katsuragi, N., N. Takizawa, and Y. Murooka.** 1987. Entire nucleotide sequence of the pullulanase gene of *Klebsiella aerogenes* W70. J. Bacteriol. **169:**2301–2306.
86. **Katz, M. E., C. L. Wright, T. S. Gartside, B. F. Cheetham, C. V. Doidge, E. K. Moses, and J. I. Rood.** 1994. Genetic organization of the duplicated *vap* region of the *Dichelobacter nodosus* genome. J. Bacteriol. **176:**2663–2669.
87. **Kaufman, M. R., J. M. Seyer, and R. K. Taylor.** 1991. Processing of TCP pilin by TcpJ typifies a common step intrinsic to a newly recognized pathway of extracellular protein secretion by gram-negative bacteria. Genes Dev. **5:**1834–1846.
88. **Kawagishi, I., V. Muller, A. W. Williams, V. M. Irikura, and R. M. Macnab.** 1992. Subdivision of flagellar region III of the *Escherichia coli* and *Salmonella typhimurium* chromosomes and identification of two additional flagellar genes. J. Gen. Microbiol. **138:**1051–1065.
89. **Kihara, M., M. Homma, K. Kutsukake, and R. M. Macnab.** 1989. Flagellar switch of *Salmonella typhimurium*: gene sequences and deduced protein sequences. J. Bacteriol. **171:**3247–3257.
90. **Kim, K., S. Lee, K. Lee, and D. Lim.** 1998. Isolation and characterization of toluene-sensitive mutants from the toluene-resistant bacterium *Pseudomonas putida* GM73. J. Bacteriol. **180:**3692–3696.
91. **Klein, J. R., B. Henrich, and R. Plapp.** 1991. Molecular analysis and nucleotide sequence of the *envCD* operon of *Escherichia coli*. Mol. Gen. Genet. **230:**230–240.
92. **Klemm, P., and G. Christiansen.** 1990. The *fimD* gene required for cell surface localization of *Escherichia coli* type 1 fimbriae. Mol. Gen. Genet. **220:**334–338.
93. **Klenk, H. P., R. A. Clayton, J. F. Tomb, O. White, K. E. Nelson, K. A. Ketchum, R. J. Dodson, M. Gwinn, E. K. Hickey, J. D. Peterson, D. L. Richardson, A. R. Kerlavage, D. E. Graham, N. C. Kyrpides, R. D. Fleischmann, J. Quackenbush, N. H. Lee, G. G. Sutton, S. Gill, E. F. Kirkness, B. A. Dougherty, K. McKenney, M. D. Adams, B. Loftus, J. C. Venter, et al.** 1997. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. Nature **390:**364–370.
94. **Kloser, A. W., M. W. Laird, and R. Misra.** 1996. *asmB*, a suppressor locus for assembly-defective OmpF mutants of *Escherichia coli*, is allelic to *envA* (*lpxC*). J. Bacteriol. **178:**5138–5143.
95. **Kobayashi, T., I. Kudo, K. Karasawa, H. Mizushima, K. Inoue, and S. Nojima.** 1985. Nucleotide sequence of the *pldB* gene and characteristics of deduced amino acid sequence of lysophospholipase L2 in *Escherichia coli*. J. Biochem. (Tokyo) **98:**1017–1025.
96. **Konz, D., S. Doekel, and M. A. Marahiel.** 1999. Molecular and biochemical characterization of the protein template controlling biosynthesis of the lipopeptide lichenysin. J. Bacteriol. **181:**133–140.
97. **Konz, D., A. Klens, K. Schorgendorfer, and M. A. Marahiel.** 1997. The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC 10716: molecular characterization of three multi-modular peptide synthetases. Chem. Biol. **4:**927–937.
98. **Kornacker, M. G., and A. P. Pugsley.** 1990. Molecular characterization of *pulA* and its product, pullulanase, a secreted enzyme of *Klebsiella pneumoniae* UNF5023. Mol. Microbiol. **4:**73–85.
99. **Kunst, F., N. Ogasawara, I. Moszer, A. M. Albertini, G. Alloni, V. Azevedo, M. G. Bertero, P. Bessieres, A. Bolotin, S. Borchert, R. Borriss, L. Boursier, A. Brans, M. Braun, S. C. Brignell, S. Bron, S. Brouillet, C. V. Bruschi, B. Caldwell, V. Capuano, N. M. Carter, S. K. Choi, J. J. Codani, I. F. Connerton, A. Danchin, et al.** 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature **390:**249–256.
100. **Kyostio, S. R., C. L. Cramer, and G. H. Lacy.** 1991. *Erwinia carotovora* subsp. *carotovora* extracellular protease: characterization and nucleotide sequence of the gene. J. Bacteriol. **173:**6537–6546.
101. **Lambert, B., L. Buysse, C. Decock, S. Jansens, C. Piens, B. Saey, J. Seurinck, K. Van Audenhove, J. Van Rie, A. Van Vliet, and M. Peferoen.** 1996. A *Bacillus thuringiensis* insecticidal crystal protein with a high activity against members of the family Noctuidae. Appl. Environ. Microbiol. **62:**80–86.
102. **Larson, T. G., and S. H. Goodgal.** 1991. Sequence and transcriptional regulation of com101A, a locus required for genetic transformation in *Haemophilus influenzae*. J. Bacteriol. **173:**4683–4691.
103. **Leclerc, M. C., and N. E. Boemare.** 1991. Plasmids and phase variation in *Xenorhabdus* spp. Appl. Environ. Microbiol. **57:**2597–2601.
104. **Le Gouill, C., D. Desmarais, and C. V. Dery.** 1993. *Saccharopolyspora hirsuta* 367 encodes clustered genes similar to ketoacyl synthase, ketoacyl reductase, acyl carrier protein, and biotin carboxyl carrier protein. Mol. Gen. Genet. **240:**146–150.
105. **Lemaire, H. G., and B. Muller-Hill.** 1986. Nucleotide sequences of the *galE* gene and the *galT* gene of *E. coli*. Nucleic Acids Res. **14:**7705–7711.
106. **Letoffe, S., P. Delepelaire, and C. Wandersman.** 1989. Characterization of a protein inhibitor of extracellular proteases produced by *Erwinia chrysanthemi*. Mol. Microbiol. **3:**79–86.
107. **Letoffe, S., P. Delepelaire, and C. Wandersman.** 1990. Protease secretion by *Erwinia chrysanthemi*: the specific secretion functions are analogous to those of *Escherichia coli* alpha-haemolysin. EMBO J. **9:**1375–1382.
108. **Levengood, S. K., and R. E. Webster.** 1989. Nucleotide sequences of the *tolA* and *tolB* genes and localization of their products, components of a multistep translocation system in *Escherichia coli*. J. Bacteriol. **171:**6600–6609.
109. **Lewis, J. P., and F. L. Macrina.** 1998. IS*195*, an insertion sequence-like element associated with protease genes in *Porphyromonas gingivalis*. Infect. Immun. **66:**3035–3042.
110. **Li, M., F. Dyda, I. Benhar, I. Pastan, and D. R. Davies.** 1996. Crystal structure of the catalytic domain of *Pseudomonas* exotoxin A complexed with a nicotinamide adenine dinucleotide analog: implications for the activation process and for ADP ribosylation. Proc. Natl. Acad. Sci. USA **93:**6902–6906.
111. **Lin, J. W., K. Y. Yu, Y. F. Chao, and S. F. Weng.** 1995. The *lumQ* gene is linked to the *lumP* gene and the *lux* operon in *Photobacterium leiognathi*. Biochem. Biophys. Res. Commun. **217:**684–695.
112. **Lin, W., K. J. Fullner, R. Clayton, J. A. Sexton, M. B. Rogers, K. E. Calia, S. B. Calderwood, C. Fraser, and J. J. Mekalanos.** 1999. Identification of a *Vibrio cholerae* RTX toxin gene cluster that is tightly linked to the cholera toxin prophage. Proc. Natl. Acad. Sci. USA **96:**1071–1076.
113. **Liu, T., Y. Lin, T. Cislo, C. A. Minetti, J. M. Baba, and T. Y. Liu.** 1991. Limunectin. A phosphocholine-binding protein from *Limulus* amebocytes with adhesion-promoting properties. J. Biol. Chem. **266:**14813–14821.
114. **Lu, D., B. Boyd, and C. A. Lingwood.** 1997. Identification of the key protein for zinc uptake in *Haemophilus influenzae*. J. Biol. Chem. **272:**29033–29038.
115. **Mahillon, J., H. A. Kirkpatrick, H. L. Kijenski, C. A. Bloch, C. K. Rode, G. F. Mayhew, D. J. Rose, G. Plunkett 3rd, V. Burland, and F. R. Blattner.** 1998. Subdivision of the *Escherichia coli* K-12 genome for sequencing: manipulation and DNA sequence of transposable elements introducing unique restriction sites. Gene **223:**47–54.
116. **Masaki, H., M. Toba, and T. Ohta.** 1985. Structure and expression of the ColE2-P9 immunity gene. Nucleic Acids Res. **13:**1623–1635.
117. **Massad, G., J. F. Fulkerson, Jr., D. C. Watson, and H. L. Mobley.** 1996. *Proteus mirabilis* ambient-temperature fimbriae: cloning and nucleotide sequence of the *aft* gene cluster. Infect. Immun. **64:**4390–4395.
118. **Matsubara, K., K. Ohnishi, and K. Kiritani.** 1992. Nucleotide sequences and characterization of *liv* genes encoding components of the high-affinity branched-amino acid transport system in *Salmonella typhimurium*. J. Biochem. (Tokyo) **112:**93–101.
119. **McClelland, M., and R. K. Wilson.** 1998. Comparison of sample sequences of the *Salmonella typhi* genome to the sequence of the complete *Escherichia coli* K-12 genome. Infect. Immun. **66:**4305–4312.
120. **McGinnes, L. W., A. Wilde, and T. G. Morrison.** 1987. Nucleotide sequence of the gene encoding the Newcastle disease virus hemagglutinin-neuraminidase protein and comparisons of paramyxovirus hemagglutinin-neuraminidase protein sequences. Virus Res. **7:**187–202.
121. **McHenney, M. A., T. J. Hosted, B. S. Dehoff, P. R. Rosteck, Jr., and R. H. Baltz.** 1998. Molecular cloning and physical mapping of the daptomycin gene cluster from *Streptomyces roseosporus*. J. Bacteriol. **180:**143–51.
122. **Merriman, T. R., M. E. Merriman, and I. L. Lamont.** 1995. Nucleotide sequence of *pvdD*, a pyoverdine biosynthetic gene from *Pseudomonas aeruginosa*: PvdD has similarity to peptide synthetases. J. Bacteriol. **177:**252–258.
123. **Michiels, T., J. C. Vanooteghem, C. Lambert de Rouvroit, B. China, A. Gustin, P. Boudry, and G. R. Cornelis.** 1991. Analysis of *virC*, an operon involved in the secretion of Yop proteins by *Yersinia enterocolitica*. J. Bacteriol. **173:**4994–5009.
124. **Miller, V. L., J. B. Bliska, and S. Falkow.** 1990. Nucleotide sequence of the *Yersinia enterocolitica ail* gene and characterization of the Ail protein product. J. Bacteriol. **172:**1062–1069.
125. **Molloy, M. P., B. R. Herbert, B. J. Walsh, M. I. Tyler, M. Traini, J. C. Sanchez, D. F. Hochstrasser, K. L. Williams, and A. A. Gooley.** 1998. Extraction of membrane proteins by differential solubilization for separation using two-dimensional gel electrophoresis. Electrophoresis **19:**837–844.
126. **Mootz, H. D., and M. A. Marahiel.** 1997. The tyrocidine biosynthesis operon of *Bacillus brevis*: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. J. Bacteriol. **179:**6843–6850.
127. **Morlon, J., R. Lloubes, S. Varenne, M. Chartier, and C. Lazdunski.** 1983. Complete nucleotide sequence of the structural gene for colicin A, a gene translated at non-uniform rate. J. Mol. Biol. **170:**271–285.
128. **Nagasawa, S., K. Ishige, and T. Mizuno.** 1993. Novel members of the two-component signal transduction genes in *Escherichia coli*. J. Biochem. (Tokyo) **114:**350–357.
129. **Naroditskaya, V., M. J. Schlosser, N. Y. Fang, and K. Lewis.** 1993. An *E. coli* gene *emrD* is involved in adaptation to low energy shock. Biochem. Biophys. Res. Commun. **196:**803–809.

130. **Nichols, W. A., S. Clegg, and M. R. Brown.** 1990. Characterization of the type 1 fimbrial subunit gene (*fimA*) of *Serratia marcescens*. Mol. Microbiol. **4:**2119–2126.
131. **Nishimura, K., T. Nakayashiki, and H. Inokuchi.** 1993. Cloning and sequencing of the *hemE* gene encoding uroporphyrinogen III decarboxylase (UPD) from *Escherichia coli* K-12. Gene **133:**109–113.
132. **Nurse, P., R. J. DiGate, K. H. Zavitz, and K. J. Marians.** 1990. Molecular cloning and DNA sequence analysis of *Escherichia coli priA*, the gene encoding the primosomal protein replication factor Y. Proc. Natl. Acad. Sci. USA **87:**4615–4619.
133. **Ogasawara, N., S. Nakai, and H. Yoshikawa.** 1994. Systematic sequencing of the 180 kilobase region of the *Bacillus subtilis* chromosome containing the replication origin. DNA Res. **1:**1–14.
134. **Olson, E. R., D. S. Dunyak, L. M. Jurss, and R. A. Poorman.** 1991. Identification and characterization of *dppA*, an *Escherichia coli* gene encoding a periplasmic dipeptide transport protein. J. Bacteriol. **173:**234–244.
135. **Oshima, T., H. Aiba, T. Baba, K. Fujita, K. Hayashi, A. Honjo, K. Ikemoto, T. Inada, T. Itoh, M. Kajihara, K. Kanai, K. Kashimoto, S. Kimura, M. Kitagawa, K. Makino, S. Masuda, T. Miki, K. Mizobuchi, H. Mori, K. Motomura, Y. Nakamura, H. Nashimoto, Y. Nishio, N. Saito, T. Horiuchi, et al.** 1996. A 718-kb DNA sequence of the *Escherichia coli* K-12 genome corresponding to the 12.7-28.0 min region on the linkage map. DNA Res. **3:**137–155.
136. **Ostrowski, J., J. Y. Wu, D. C. Rueger, B. E. Miller, L. M. Siegel, and N. M. Kredich.** 1989. Characterization of the *cysJIH* regions of *Salmonella typhimurium* and *Escherichia coli* B. DNA sequences of *cysI* and *cysH* and a model for the siroheme-Fe4S4 active center of sulfite reductase hemoprotein based on amino acid homology with spinach nitrite reductase. J. Biol. Chem. **264:**15726–15737.
137. **Oswald, E., M. Sugai, A. Labigne, H. C. Wu, C. Fiorentini, P. Boquet, and A. D. O'Brien.** 1994. Cytotoxic necrotizing factor type 2 produced by virulent *Escherichia coli* modifies the small GTP-binding proteins Rho involved in assembly of actin stress fibers. Proc. Natl. Acad. Sci. USA **91:**3814–3818.
138. **Peakman, T., J. Crouzet, J. F. Mayaux, S. Busby, S. Mohan, N. Harborne, J. Wootton, R. Nicolson, and J. Cole.** 1990. Nucleotide sequence, organisation and structural analysis of the products of genes in the nirB-cysG region of the *Escherichia coli* K-12 chromosome. Eur. J. Biochem. **191:**315–323.
139. **Pelludat, C., A. Rakin, C. A. Jacobi, S. Schubert, and J. Heesemann.** 1998. The yersiniabactin biosynthetic gene cluster of *Yersinia enterocolitica*: organization and siderophore-dependent regulation. J. Bacteriol. **180:**538–546.
140. **Perry, R. D., and J. D. Fetherston.** 1997. *Yersinia pestis*—etiologic agent of plague. Clin. Microbiol. Rev. **10:**35–66.
141. **Poole, K., E. Schiebel, and V. Braun.** 1988. Molecular characterization of the hemolysin determinant of *Serratia marcescens*. J. Bacteriol. **170:**3177–3188.
142. **Popham, D. L., and P. Setlow.** 1993. Cloning, nucleotide sequence, and regulation of the *Bacillus subtilis pbpE* operon, which codes for penicillin-binding protein 4* and an apparent amino acid racemase. J. Bacteriol. **175:**2917–2925.
143. **Pospiech, A., B. Cluzel, J. Bietenhader, and T. Schupp.** 1995. A new *Myxococcus xanthus* gene cluster for the biosynthesis of the antibiotic saframycin Mx1 encoding a peptide synthetase. Microbiology **141:**1793–1803.
144. **Price, C., J. Lingner, T. A. Bickle, K. Firman, and S. W. Glover.** 1989. Basis for changes in DNA recognition by the *Eco*R124 and *Eco*R124/3 type I DNA restriction and modification enzymes. J. Mol. Biol. **205:**115–125.
145. **Raina, S., D. Missiakas, L. Baird, S. Kumar, and C. Georgopoulos.** 1993. Identification and transcriptional analysis of the *Escherichia coli htrE* operon which is homologous to *pap* and related pilin operons. J. Bacteriol. **175:**5009–5021.
146. **Rakin, A., E. Saken, D. Harmsen, and J. Heesemann.** 1994. The pesticin receptor of *Yersinia enterocolitica*: a novel virulence factor with dual function. Mol. Microbiol. **13:**253–263.
147. **Reich, K. A., and G. K. Schoolnik.** 1996. Halovibrin, secreted from the light organ symbiont *Vibrio fischeri*, is a member of a new class of ADP-ribosyltransferases. J. Bacteriol. **178:**209–215.
148. **Relman, D. A., M. Domenighini, E. Tuomanen, R. Rappuoli, and S. Falkow.** 1989. Filamentous hemagglutinin of *Bordetella pertussis*: nucleotide sequence and crucial role in adherence. Proc. Natl. Acad. Sci. USA **86:**2637–2641.
149. **Roberts, R. C., and D. R. Helinski.** 1992. Definition of a minimal plasmid stabilization system from the broad-host-range plasmid RK2. J. Bacteriol. **174:**8119–8132.
150. **Roman, S. J., B. B. Frantz, and P. Matsumura.** 1993. Gene sequence, overproduction, purification and determination of the wild-type level of the *Escherichia coli* flagellar switch protein FliG. Gene **133:**103–108.
151. **Rosqvist, R., M. Skurnik, and H. Wolf-Watz.** 1988. Increased virulence of *Yersinia pseudotuberculosis* by two independent mutations. Nature **334:**522–524.
152. **Saito, F., K. Hori, M. Kanda, T. Kurotsu, and Y. Saito.** 1994. Entire nucleotide sequence for *Bacillus brevis* Nagano Grs2 gene encoding gramicidin S synthetase 2: a multifunctional peptide synthetase. J. Biochem. (Tokyo) **116:**357–367.
153. **Sancar, G. B., A. Sancar, and W. D. Rupp.** 1984. Sequence of the *E. coli uvrC* gene and protein. Nucleic Acids Res. **12:**4593–4608.
154. **Sauerborn, M., and C. von Eichel-Streiber.** 1990. Nucleotide sequence of *Clostridium difficile* toxin A. Nucleic Acids Res. **18:**1629–1630.
155. **Schaffer, S., K. Hantke, and V. Braun.** 1985. Nucleotide sequence of the iron regulatory gene *fur*. Mol. Gen. Genet. **200:**110–113.
156. **Schmitt, M. P., and R. K. Holmes.** 1994. Cloning, sequence, and footprint analysis of two promoter/operators from *Corynebacterium diphtheriae* that are regulated by the diphtheria toxin repressor (DtxR) and iron. J. Bacteriol. **176:**1141–1149.
157. **Seoane, A., and J. M. Garcia Lobo.** 1991. Nucleotide sequence of a new class A beta-lactamase gene from the chromosome of *Yersinia enterocolitica*: implications for the evolution of class A beta-lactamases. Mol. Gen. Genet. **228:**215–220.
158. **Skurnik, M., R. Venho, P. Toivanen, and A. al-Hendy.** 1995. A novel locus of *Yersinia enterocolitica* serotype O:3 involved in lipopolysaccharide outer core biosynthesis. Mol. Microbiol. **17:**575–594.
159. **Staudenmaier, H., B. Van Hove, Z. Yaraghi, and V. Braun.** 1989. Nucleotide sequences of the *fecBCDE* genes and locations of the proteins suggest a periplasmic-binding-protein-dependent transport mechanism for iron(III) dicitrate in *Escherichia coli*. J. Bacteriol. **171:**2626–2633.
160. **Stein, D. C., R. Chien, and H. S. Seifert.** 1992. Construction of a *Neisseria gonorrhoeae* MS11 derivative deficient in *Ngo*MI restriction and modification. J. Bacteriol. **174:**4899–4906.
161. **Stevenson, G., K. Andrianopoulos, M. Hobbs, and P. R. Reeves.** 1996. Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. J. Bacteriol. **178:**4885–4893.
162. **Stintzi, A., Z. Johnson, M. Stonehouse, U. Ochsner, J. M. Meyer, M. L. Vasil, and K. Poole.** 1999. The *pvc* gene cluster of *Pseudomonas aeruginosa*: role in synthesis of the pyoverdine chromophore and regulation by PtxR and PvdS. J. Bacteriol. **181:**4118–4124.
163. **Stirling, D. A., C. S. Hulton, L. Waddell, S. F. Park, G. S. Stewart, I. R. Booth, and C. F. Higgins.** 1989. Molecular characterization of the *proU* loci of *Salmonella typhimurium* and *Escherichia coli* encoding osmoregulated glycine betaine transport systems. Mol. Microbiol. **3:**1025–1038.
164. **Stock, A., T. Chen, D. Welsh, and J. Stock.** 1988. CheA protein, a central regulator of bacterial chemotaxis, belongs to a family of proteins that control gene expression in response to changing environmental conditions. Proc. Natl. Acad. Sci. USA **85:**1403–1407.
165. **Stojiljkovic, I., and K. Hantke.** 1994. Transport of haemin across the cytoplasmic membrane through a haemin-specific periplasmic binding-protein-dependent transport system in *Yersinia enterocolitica*. Mol. Microbiol. **13:**719–732.
166. **Stone, B. J., and V. L. Miller.** 1995. *Salmonella enteritidis* has a homologue of *tolC* that is required for virulence in BALB/c mice. Mol. Microbiol. **17:**701–712.
167. **Ta, D. T., and L. E. Vickery.** 1992. Cloning, sequencing, and overexpression of a [2Fe-2S] ferredoxin gene from *Escherichia coli*. J. Biol. Chem. **267:**11120–11125.
168. **Tabatabai, N., and S. Forst.** 1995. Molecular analysis of the two-component genes, *ompR* and *envZ*, in the symbiotic bacterium *Xenorhabdus nematophilus*. Mol. Microbiol. **17:**643–652.
169. **Taron, C. H., E. M. Van Cott, G. G. Wilson, L. S. Moran, B. E. Slatko, L. J. Hornstra, J. S. Benner, R. B. Kucera, and E. P. Guthrie.** 1995. Cloning and expression of the *Nae*I restriction endonuclease-encoding gene and sequence analysis of the *Nae*I restriction-modification system. Gene **155:**19–25.
170. **Tercero, J. A., J. C. Espinosa, R. A. Lacalle, and A. Jimenz.** 1996. The biosynthetic pathway of the aminonucleoside antibiotic puromycin, as deduced from the molecular analysis of the *pur* cluster of *Streptomyces alboniger*. J. Biol. Chem. **271:**1579–1590.
171. **Tews, I., R. Vincentelli, and C. E. Vorgias.** 1996. *N*-Acetylglucosaminidase (chitobiase) from *Serratia marcescens*: gene sequence, and protein production and purification in *Escherichia coli*. Gene **170:**63–67.
172. **Tobe, T., C. Sasakawa, N. Okada, Y. Honma, and M. Yoshikawa.** 1992. *vacB*, a novel chromosomal gene required for expression of virulence genes on the large plasmid of *Shigella flexneri*. J. Bacteriol. **174:**6359–6367.
173. **Tomb, J. F., H. el-Hajj, and H. O. Smith.** 1991. Nucleotide sequence of a cluster of genes involved in the transformation of *Haemophilus influenzae* Rd. Gene **104:**1–10.
174. **Uphoff, T. S., and R. A. Welch.** 1990. Nucleotide sequencing of the *Proteus mirabilis* calcium-independent hemolysin genes (*hpmA* and *hpmB*) reveals sequence similarity with the *Serratia marcescens* hemolysin genes (*shlA* and *shlB*). J. Bacteriol. **172:**1206–1216.
175. **van Buul, C. P., and P. H. van Knippenberg.** 1985. Nucleotide sequence of the *ksgA* gene of *Escherichia coli*: comparison of methyltransferases effecting dimethylation of adenosine in ribosomal RNA. Gene **38:**65–72.
176. **van der Ploeg, J. R., M. A. Weiss, E. Saller, H. Nashimoto, N. Saito, M. A. Kertesz, and T. Leisinger.** 1996. Identification of sulfate starvation-regu-

lated genes in *Escherichia coli*: a gene cluster involved in the utilization of taurine as a sulfur source. J. Bacteriol. **178:**5438–5446.
177. **van Montfort, R. L., T. Pijning, K. H. Kalk, I. Hangyi, M. L. Kouwijzer, G. T. Robillard, and B. W. Dijkstra.** 1998. The structure of the *Escherichia coli* phosphotransferase IIA mannitol reveals a novel fold with two conformations of the active site. Structure **6:**377–388.
178. **Viitanen, A. M., P. Toivanen, and M. Skurnik.** 1990. The *lcrE* gene is part of an operon in the *lcr* region of *Yersinia enterocolitica* O:3. J. Bacteriol. **172:**3152–3162.
179. **von Wilcken-Bergmann, B., and B. Muller-Hill.** 1982. Sequence of *galR* gene indicates a common evolutionary origin of *lac* and *gal* repressor in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **79:**2427–2431.
180. **Vuorio, R., T. Harkonen, M. Tolvanen, and M. Vaara.** 1994. The novel hexapeptide motif found in the acyltransferases LpxA and LpxD of lipid A biosynthesis is conserved in various bacteria. FEBS Lett. **337:**289–292.
181. **Wang, H., and B. C. Dowds.** 1993. Phase variation in *Xenorhabdus luminescens*: cloning and sequencing of the lipase gene and analysis of its expression in primary and secondary phases of the bacterium. J. Bacteriol. **175:** 1665–1673.
182. **Welch, R. A., M. E. Bauer, A. D. Kent, J. A. Leeds, M. Moayeri, L. B. Regassa, and D. L. Swenson.** 1995. Battling against host phagocytes: the wherefore of the RTX family of toxins? Infect. Agents Dis. **4:**254–272.
183. **Whitchurch, C. B., and J. S. Mattick.** 1994. *Escherichia coli* contains a set of genes homologous to those involved in protein secretion, DNA uptake and the assembly of type-4 fimbriae in other bacteria. Gene **150:**9–15.
184. **Williams, S. G., L. T. Varcoe, S. R. Attridge, and P. A. Manning.** 1996. *Vibrio cholerae* Hcp, a secreted protein coregulated with HlyA. Infect. Immun. **64:**283–289.
185. **Wipat, A., S. C. Brignell, B. J. Guy, M. Rose, P. T. Emmerson, and C. R. Harwood.** 1998. The *yvsA-yvqA* (293 degrees-289 degrees) region of the *Bacillus subtilis* chromosome containing genes involved in metal ion uptake and a putative sigma factor. Microbiology **144:**1593–1600.
186. **Xu, K., and T. Elliott.** 1993. An oxygen-dependent coproporphyrinogen oxidase encoded by the *hemF* gene of *Salmonella typhimurium*. J. Bacteriol. **175:**4990–4999.
187. **Yamagata, H., K. Nakamura, and M. Inouye.** 1981. Comparison of the lipoprotein gene among the Enterobacteriaceae. DNA sequence of *Erwinia amylovora* lipoprotein gene. J. Biol. Chem. **256:**2194–2198.
188. **Yamamoto, H., S. Uchiyama, and J. Sekiguchi.** 1996. The *Bacillus subtilis* chromosome region near 78 degrees contains the genes encoding a new two-component system, three ABC transporters and a lipase. Gene **181:** 147–151.
189. **Yamamoto, Y., H. Aiba, T. Baba, K. Hayashi, T. Inada, K. Isono, T. Itoh, S. Kimura, M. Kitagawa, K. Makino, T. Miki, N. Mitsuhashi, K. Mizobuchi, H. Mori, S. Nakade, Y. Nakamura, H. Nashimoto, T. Oshima, S. Oyama, N. Saito, G. Sampei, Y. Satoh, S. Sivasundaram, H. Tagami, T. Horiuchi, et al.** 1997. Construction of a contiguous 874-kb sequence of the *Escherichia coli* K12 genome corresponding to 50.0-68.8 min on the linkage map and analysis of its sequence features. DNA Res. **4:**91–113.
190. **Yang, Y., J. J. Merriam, J. P. Mueller, and R. R. Isberg.** 1996. The *psa* locus is responsible for thermoinducible binding of *Yersinia pseudotuberculosis* to cultured cells. Infect. Immun. **64:**2483–2489.
191. **Yoshida, K., K. Shindo, H. Sano, S. Seki, M. Fujimura, N. Yanai, Y. Miwa, and Y. Fujita.** 1996. Sequencing of a 65 kb region of the *Bacillus subtilis* genome containing the *lic* and *cel* loci, and creation of a 177 kb contig covering the *gnt-sacXY* region. Microbiology **142:**3113–3123.
192. **Zavitz, K. H., R. J. DiGate, and K. J. Marians.** 1991. The priB and priC replication proteins of *Escherichia coli*. Genes, DNA sequence, overexpression, and purification. J. Biol. Chem. **266:**13988–13995.
193. **Zenno, S., and K. Saigo.** 1994. Identification of the genes encoding NAD(P)H-flavin oxidoreductases that are similar in sequence to *Escherichia coli* Fre in four species of luminous bacteria: *Photorhabdus luminescens*, *Vibrio fischeri*, *Vibrio harveyi*, and *Vibrio orientalis*. J. Bacteriol. **176:** 3544–3551.
194. **Zulty, J. J., and G. J. Barcak.** 1995. Identification of a DNA transformation gene required for com101A+ expression and supertransformer phenotype in *Haemophilus influenzae*. Proc. Natl. Acad. Sci. USA **92:**3616–3620.